

P30530 homo sapien
Q00993 mus musculu

ALIGNMENTS

44	495.5	10.1	887	1	UFO_HUMAN	P30530	homo sapien
45	. 493	10.1	888	1	UFO_MOUSE	Q00993	mus musculus

1627.298 Million cell updates/sec

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	4895	99.8	913	1	EDD1_HUMAN	Q08345 homo sapien
2	4584.5	93.5	910	1	EDD1_RAT	Q63474 rattus norv
3	4568	93.2	911	1	EDD1_MOUSE	Q03146 mus musculu
4	2415	49.3	855	1	TRK3_HUMAN	Q16832 homo sapien
5	2413	49.2	854	1	TRK3_MOUSE	Q62371 mus musculu
6	661	13.5	818	1	TRKB_CHICK	Q91987 gallus gall
7	659.5	13.5	821	1	TRKB_MOUSE	Q15209 mus musculu
8	659.5	13.5	821	1	TRKB_RAT	Q63604 rattus norv
9	658	13.3	822	1	TRKB_HUMAN	Q16620 homo sapien
10	654.5	13.3	796	1	TRKA_HUMAN	Q16529 homo sapien
11	653.5	13.3	799	1	TRKA_RAT	P35739 rattus norv
12	644.5	13.1	778	1	TRKA_CHICK	P35739 rattus norv
13	642.5	13.1	825	1	TRKC_PIG	Q91009 gallus gall
14	633.5	12.9	839	1	TRKC_HUMAN	P24786 mus scrofa
15	614	12.5	852	1	TRKC_CHICK	Q16288 homo sapien
16	611.5	12.5	864	1	TRKC_RAT	Q19044 gallus gall
17	601.5	12.3	605	1	RTK2_GEOCY	Q91351 rattus norv
18	544.5	11.1	402	1	KROS_AVISU	P42159 geodia cydo
19	544.5	11.1	507	1	KROS_CHICK	P00529 avian sarco
20	539.5	11.0	1363	1	IUPR_BRALA	P08941 gallus gall
21	535	10.9	2347	1	KROS_HUMAN	Q02466 branchiost
22	534.5	10.9	1382	1	INSR_HUMAN	P08922 homo sapien
23	530.5	10.8	1372	1	INSR_MOUSE	P06213 homo sapien
24	528	10.8	1520	1	ABL_DROME	P15208 mus musculu
25	527.5	10.8	1383	1	INSR_RAT	P00522 drosophila
26	525	10.7	1300	1	INSR_CAVPO	P15127 rattus norv
27	523	10.7	746	1	ABL_MVAB	P14617 cavia porce
28	523	10.7	1123	1	ABL_MOUSE	P00521 abelson mur
29	521.5	10.6	1379	1	MET_MOUSE	P00520 mus musculu
30	521	10.6	1268	1	IRR_HUMAN	P16056 mus musculu
31	520	10.6	1130	1	ABL1_HUMAN	P14616 homo sapien
32	520	10.6	2146	1	INSR_DROME	P00519 homo sapien
33	518	10.6	1390	1	MET_HUMAN	P09208 drosophila
34	517	10.5	2594	1	7LES_DROYT	P08581 homo sapien
35	517	10.5	439	1	ABL_FSVHY	P20806 drosophila
36	511	10.4	1182	1	ABL2_HUMAN	P10447 feline sarc
37	511	10.4	1367	1	IGIR_HUMAN	P42684 homo sapien
38	509	10.4	640	1	IGIR_BOVIN	P08069 homo sapien
39	508.5	10.4	2554	1	7LES_DROME	Q05688 bos taurus
40	505.5	10.3	802	1	FGR4_HUMAN	P13368 drosophila
41	503	10.3	1114	1	RGT_HUMAN	P22455 homo sapien
42	503	10.3	1370	1	IGIR_RAT	P07949 homo sapien
43	497	10.1	1607	1	MIPR_LYMSI	P24062 rattus norv
						Q25410 lynnaea sta

RESULT 1

EDDI_HUMAN
ID EDDI_HUMAN STANDARD; PRT; 913 AA.
Q0345; Q14196; Q16562;
AC 01-FEB-1995 (REL. 31, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DE EPIHELIAL DISCOIDIN DOMAIN RECEPTOR 1 PRECURSOR (EC 2.7.1.112)
DE (TYROSINE-PROTEIN KINASE CAK) (CELL ADHESION KINASE) (TYROSINE KINASE
DE DDR) (DISCOIDIN RECEPTOR TYROSINE KINASE) (TRK E) (PROTEIN-TYROSINE
DE KINASE RTK 6).
GN EDDR1 OR CAK OR DDR1 OR TRKE OR RTK6.
OS HOMO SAPIENS (HUMAN).
OC EUMAROTATA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
[1]
SEQUENCE FROM N.A.
RPP TISSUE-FETAL LIVER;
RPP MEDLINE; 94134417.
RRC PEREZ J.L., SHEN X., FINKERNAGEL S., SCIORRA L., JENKINS N.A.,
RRA GILBERT D.J., COPELAND N.G., WONG T.W.;
RRT "Identification and chromosomal mapping of a receptor tyrosine kinase
RRT with a putative phospholipid binding sequence in its ectodomain.";
RRL ONCOGENE 9:211-219(1994).
[2]
SEQUENCE FROM N.A.
RPP TISSUE-PLACENTA;
RPP MEDLINE; 93296201.
RRC JOHNSON J.D., EDMAN J.C., RUTTER W.J.,
RRA "A receptor tyrosine kinase found in breast carcinoma cells has an
RRT extracellular discoidin I-like domain";
RRT PROC. NATL. ACAD. SCI. U.S.A. 90:5677-5681(1993).
[3]
SEQUENCE FROM N.A.
RPP MEDLINE; 97131588.
RRA SAKUMA S., TADA M., SAYA H., SAWAMURA Y., SHINOHE Y., ABE H.;
RRT "Receptor protein tyrosine kinase DDR is up-regulated by p53
RRT protein";
RRL FEBS LETT. 398:165-169(1996).
[4]
SEQUENCE FROM N.A.
RPP MEDLINE; 96389017.
RRA PLAYFORD M.P., BUTLER R.J., WANG X.C., KATSO R.M., COOKE I.E.,
RRT GANESAN T.S.;
RRT "The genomic structure of discoidin receptor tyrosine kinase";
RRL GENOME RES. 6:620-627(1996).
[5]
SEQUENCE FROM N.A. (SHORT FORM).
RPP TISSUE=LUNG;
RPP MEDLINE; 96204002.
RRC PEREZ J.L., JING S.Q., WONG T.W.;
RRA "Identification of two isoforms of the Cak receptor kinase that are
RRT coexpressed in breast tumor cell lines";
RRL ONCOGENE 12:1469-1477(1996).
[6]
SEQUENCE FROM N.A. (SHORT FORM).
RPP TISSUE-BRAIN, AND KERATINOGENES;
RPP MEDLINE; 94043285.
RRA DI MARCO E., CUTULI N., GUERRA L., CANCEDDA R., DE LUCA M.;
RRT "Molecular cloning of trkE, a novel trk-related putative tyrosine
RRT kinase receptor isolated from normal human keratinocytes and widely
RRT expressed by normal human tissues";
RRL J. BIOL. CHEM. 268:24290-24295(1993).
[7]
SEQUENCE FROM N.A. (SHORT FORM).
RPP TISSUE=OVARY.
RPP MEDLINE; 94043285.

RX MEDLINE; 95151638.
 RA LAVAL S., BUTLER R., SHELLING A.N., HANBY A.M., POULSON R.,
 RA GANESAN T.S.;
 RT "Isolation and characterization of an epithelial-specific receptor
 RL tyrosine kinase from an ovarian cancer cell line."
 RL CELL GROWTH DIFFER. 5:1173-1183(1994).
 CC -!- FUNCTION: MAY BE INVOLVED IN CELL-CELL INTERACTIONS AND
 CC RECOGNITION.
 CC -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE - ADP +
 CC PROTEIN TYROSINE PHOSPHATE.
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -!- TISSUE SPECIFICITY: EXPRESSED AT LOW LEVELS IN MOST ADULT
 CC TISSUES AND IS HIGHEST IN THE BRAIN AND LUNG. ABUNDANT IN
 CC BREAST CARCINOMA CELL LINES.
 CC -!- DOMAIN: THE GLY/PRO-RICH DOMAINS MAY BE REQUIRED FOR AN UNUSUAL
 CC GEOMETRY OF INTERACTION WITH LIGAND OR SUBSTRATES.
 CC -!- ALTERNATIVE PRODUCTS: TWO ISOFORMS ARE PRODUCED BY ALTERNATIVE
 CC SPLICING. THE LONG FORM (CALLED CAK I AND SHOWN HERE) AND THE
 CC SHORT FORM (CALLED CAK II) ONLY DIFFER BY THE ABSENCE OF A 37
 CC RESIDUES SEGMENT.
 CC -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
 CC DOMAIN.
 CC -!- SIMILARITY: CONTAINS 1 F5/8 TYPE C DOMAIN.
 CC -!- SIMILARITY: BELONGS TO THE INSULIN RECEPTOR FAMILY OF TYROSINE-
 CC PROTEIN KINASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L11315; G403387; -
 DR EMBL; L20817; G306475; -
 DR EMBL; U48705; G1754536; -
 DR EMBL; X98208; E251231; JOINED.
 DR EMBL; X99023; E251231; JOINED.
 DR EMBL; X99024; E251231; JOINED.
 DR EMBL; X99025; E251231; JOINED.
 DR EMBL; X99026; E251231; JOINED.
 DR EMBL; X99027; E251231; JOINED.
 DR EMBL; X99028; E251231; JOINED.
 DR EMBL; X99029; E251231; JOINED.
 DR EMBL; X99030; E251231; JOINED.
 DR EMBL; X99031; E251231; JOINED.
 DR EMBL; X99032; E251231; JOINED.
 DR EMBL; X99033; E251231; JOINED.
 DR EMBL; X99034; E251231; JOINED.
 DR EMBL; L57508; G1160925; -
 DR EMBL; X74979; G400463; -
 DR EMBL; Z29093; G732800; -
 DR MIM; 600452; -
 DR MIM; 600408; -
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.
 DR PROSITE; PS01285; FA58C_1; 1.
 DR PROSITE; PS01286; FA58C_2; 1.
 DR PFAM; PF00069; pkinase; 1.
 DR PFAM; PF00754; F5_F8_type_C; 1.
 DR HSP; P11362; IFC1.
 DR TRANSFERASE; TYROSINE-PROTEIN KINASE; GLYCOPROTEIN; SIGNAL;
 KW PHOSPHORYLATION; TRANSMEMBRANE; RECEPTOR; ATP-BINDING;
 KW ALTERNATIVE SPLICING.
 FT SIGNAL 18
 FT CHAIN 19
 FT DOMAIN 19
 FT TRANSMEM 417
 FT DOMAIN 444
 FT TRANSMEM 444
 FT DOMAIN 31
 FT DOMAIN 185
 F5/8 TYPE C (PHOSPHOLIPID-BINDING,

FT DOMAIN 377 415
 FT DOMAIN 476 601
 FT DOMAIN 610 905
 FT NP_BIND 616 624
 FT BINDING 655 655
 FT ACT_SITE 766 766
 FT DISULFID 31 185
 FT MOD_RES 513 513
 FT MOD_RES 792 792
 FT MOD_RES 796 796
 FT MOD_RES 797 797
 FT CARBOHYD 211 211
 FT CARBOHYD 260 260
 FT CARBOHYD 371 371
 FT CARBOHYD 394 394
 FT VARSPLIC 506 542
 FT CONFLICT 94 94
 FT CONFLICT 833 833
 FT CONFLICT 847 867
 FT CONFLICT 847 867
 SQ SEQUENCE 913 AA; 101127 MW; 16071364 CRC32;
 Query Match 99.8%; Score 4895; DB 1; Length 913;
 Best Local Similarity 99.9%; Pred. No. 1.9e-242;
 Matches 912; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MGPEALSLILLVLSVSGADAMKGFDPKACRYALGMQDRTPDSDISASSWSDDTAAR 60
 DB 1 MGPEALSLILLVLSVSGADAMKGFDPKACRYALGMQDRTPDSDISASSWSDDTAAR 60
 QY 61 HSRLESSDGDGAWCPAGSVFPKEEYLQVDLQRLHLVALVGTQGRHAGGLCKEFSRSL 120
 DB 61 HSRLESSDGDGAWCPAGSVFPKEEYLQVDLQRLHLVALVGTQGRHAGGLCKEFSRSL 120
 QY 121 RYSDRGRRMGMKDRWGQEVISGNEDEPGVYKDLGPPMVARLVRFPRADRVMSVCLRV 180
 DB 121 RYSDRGRRMGMKDRWGQEVISGNEDEPGVYKDLGPPMVARLVRFPRADRVMSVCLRV 180
 QY 181 ELYGCLWRDGLLSYTPVGTMYLSEAVYVNDSTVDGTVGLQVGLGLQADGVVGLDD 240
 DB 181 ELYGCLWRDGLLSYTPVGTMYLSEAVYVNDSTVDGTVGLQVGLGLQADGVVGLDD 240
 QY 241 FRKSQELRWPCYDYVGNWSNHSFSSGYVEMEFEDRLRAFOAMQVHCNMHTLGLARPG 300
 DB 241 FRKSQELRWPCYDYVGNWSNHSFSSGYVEMEFEDRLRAFOAMQVHCNMHTLGLARPG 300
 QY 301 VECRRRGPMAMWEGEPHMRHNLGNLGDPRARAVSVPLGGVRVARELQCRFLFAGPWL 360
 DB 301 VECRRRGPMAMWEGEPHMRHNLGNLGDPRARAVSVPLGGVRVARELQCRFLFAGPWL 360
 QY 361 EISFTSDVYNNSSPALGTFPPAPWMPGPPPTNFSSLELEPRGQPVAKPEGSP 420
 DB 361 EISFTSDVYNNSSPALGTFPPAPWMPGPPPTNFSSLELEPRGQPVAKPEGSP 420
 QY 421 GCLVAIILLIILALMLWRLHWRLLSKAERVLLEELTVHLSVPGDTILINNRPGPRE 480
 DB 421 GCLVAIILLIILALMLWRLHWRLLSKAERVLLEELTVHLSVPGDTILINNRPGPRE 480
 QY 481 PPHYQEPGRGNPHSAPCVNGSALLSNPAYRLLATYARPPRGPGPTPAWAKPNT 540
 DB 481 PPHYQEPGRGNPHSAPCVNGSALLSNPAYRLLATYARPPRGPGPTPAWAKPNT 540
 QY 541 QAYSGDYMEPEKPGAPLLPPPPQNSVPHYAEADIVTLQVGTGGNTYAVPALPPGAVGDP 600
 DB 541 QAYSGDYMEPEKPGAPLLPPPPQNSVPHYAEADIVTLQVGTGGNTYAVPALPPGAVGDP 600
 QY 601 PRVDFPRSLRKEKLGEGQFGEVHLCVDSQDLVSLDFPLNVRKGHPLLVAVKILRPD 660
 DB 601 PRVDFPRSLRKEKLGEGQFGEVHLCVDSQDLVSLDFPLNVRKGHPLLVAVKILRPD 660
 QY 661 ATKNAERDNLKVKIMSLKDPNIIRLLGVCVQDDPDLCLMTIDYMGENDLQNFLSAHOLED 720

Db 661 ATNNARNDLFKEIKMSRLKPNILRLGVVODDPLCMITDMENGDDNQFLSAHQED 720
 QY 721 KAAEGAPGDGQAAGPTISYPMMLHVAQAQASGMRYLATLNFVHRDLATRNCLVGENFTI 780
 Db 721 KAAEGAPGDGQAAGPTISYPMMLHVAQAQASGMRYLATLNFVHRDLATRNCLVGENFTI 780
 QY 781 KIADFGMSRLNLYAGDYRVOGRAVLPIRMAWECILMGKFTTASDVWAFGVTLWEVLMC 840
 Db 781 KIADFGMSRLNLYAGDYRVOGRAVLPIRMAWECILMGKFTTASDVWAFGVTLWEVLMC 840
 QY 841 RAOPFGOLTDEQVIENAGFEFRDQGRQVLSRPPACPOGLYELMLRCWSRESEQRPPFSQ 900
 Db 841 RAOPFGOLTDEQVIENAGFEFRDQGRQVLSRPPACPOGLYELMLRCWSRESEQRPPFSQ 900
 QY 901 LHRFLAEDALNTV 913
 Db 901 LHRFLAEDALNTV 913
 RESULT 2
 EDDI_RAT
 ID EDDI_RAT STANDARD; PRT; 910 AA.
 AC O63474;
 DT 01-NOV-1997 (REL. 35, CREATED)
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE EPITHELIAL DISCOIDIN DOMAIN RECEPTOR 1 PRECURSOR (BC 2.7.1.112)
 DE (TYROSINE-PROTEIN KINASE CAK) (CELL ADHESION KINASE) (TYROSINE KINASE
 DE DDK) (DISCOIDIN RECEPTOR TYROSINE KINASE) (PROTEIN-TYROSINE KINASE
 DE PK-3)
 GN EDDRI OR PK3.
 OS RATTUS NORVEGICUS (RAT).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SPRAGUE-DAWLEY; TISSUE-BRAIN;
 RX MEDLINE; 94173920.
 RA SANCHEZ M.B., TAPLEY P., SAINI S.S., HE B., PULIDO D., BARBACID M.;
 RT Multiple tyrosine protein kinases in rat hippocampal neurons;
 RT Isolation of ptk-3, a receptor expressed in proliferative zones of
 RT the developing brain";
 RL PROC. NATL. ACAD. SCI. U.S.A. 91:1819-1823(1994).
 CC !- FUNCTION: MAY BE INVOLVED IN CELL-CELL INTERACTIONS AND
 CC RECOGNITION (BY SIMILARITY).
 CC !- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
 CC PROTEIN TYROSINE PHOSPHATE.
 CC !- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC !- TISSUE SPECIFICITY: VARIOUS EMBRYONIC AND ADULT TISSUES; ALSO
 CC PROLIFERATIVE ZONES OF THE DEVELOPING BRAIN; HIPPOCAMPAL NEURONS.
 CC !- SIMILARITY: WITH THE CONSERVED CATALYTIC DOMAINS OF TYR-PROTEIN
 CC KINASES.
 CC !- SIMILARITY: CONTAINS 1 F5/8 TYPE C DOMAIN.
 CC !- SIMILARITY: BELONGS TO THE INSULIN RECEPTOR FAMILY OF TYROSINE-
 CC PROTEIN KINASES.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; L26525; G432481; .
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.
 DR PROSITE; PS01285; FA58C_1; 1.
 DR PROSITE; PS01286; FA58C_2; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PFAM; PF00069; pkinase; 1.
 DR PFAM; PF00754; F5_F8_type_C; 1.

DR HSP; P11362; IFGI.
 KW TRANSFERASE; THROSKINE-PROTEIN KINASE; GLYCOPROTEIN; SIGNAL;
 FT PHOSPHORYLATION; TRANSMEMBRANE; RECEPTOR; ATP-BINDING.
 FT SIGNAL 1 19
 FT CHAIN 20 910
 FT DOMAIN 20 413
 FT TRANSMEM 414 440
 FT DOMAIN 441 910
 FT DOMAIN 32 186
 FT DOMAIN 378 412
 FT DOMAIN 473 598
 FT DOMAIN 607 902
 FT NP_BIND 613 621
 FT BINDING 652 621
 FT ACT_SITE 763 763
 FT DISULFID 32 186
 FT MOD_RES 510 510
 FT MOD_RES 789 789
 FT MOD_RES 793 793
 FT MOD_RES 794 794
 FT CARBOHYD 212 212
 FT CARBOHYD 261 261
 FT CARBOHYD 371 371
 FT CARBOHYD 391 391
 SQ SEQUENCE 910 AA; 101164 MW; D69094A9 CRC32;
 Query Match 93.5%; Score 4584.5; DB 1; Length 910;
 Best Local Similarity 93.5%; Pred. No. 1.2e-226;
 Matches 855; Conservative 20; Mismatches 34; Indels 5; Gaps 2;
 QY 1 MGPEALSS-LLLLLVASGDADMKGHFDPKACRYALGMQDRTIPDSDISASSWSSTAA 59
 Db 1 MGTCTLSLLLLLLLVITIGDADMKGHFDPKACRYALGMQDRTIPDSDISASSWSSTAA 60
 QY 60 RHRSLSSDGDGAWCPAGSVFPKEEYLVQDLRLHLVALVGTQGRHAGGLGKFEFSRYSR 119
 Db 61 RHRSLSSDGDGAWCPAGSVFPKEEYLVQDLRLHLVALVGTQGRHAGGLGKFEFSRYSR 120
 QY 120 LRYSDGRRRWGWKDRMGQEVISGNEDEPGVWLKDLGPPMVARLVRFYPRADRVMSVCLR 179
 Db 121 LRYSDGRRRWGWKDRMGQEVISGNEDEPGVWLKDLGPPMVARLVRFYPRADRVMSVCLR 180
 QY 180 VELYGCLWRDGLLSYAPVGTMTYLSAVYLVNDSTVDGHTVGLYGGGLQGLADGVVGLD 239
 Db 181 VELYGCLWRDGLLSYAPVGTMTYLSAVYLVNDSTVDGHTVGLYGGGLQGLADGVVGLD 240
 QY 240 DFRKSQELRYWPGYDYVGVNSHSFSSGYVEMEFEDRLRAFAQAMQVHCNNMHTLGARLPG 299
 Db 241 DFRKSQELRYWPGYDYVGVNSHSFSSGYVEMEFEDRLRAFAQAMQVHCNNMHTLGARLPG 300
 QY 300 GVECRFRGRGAMAWEGEPHRLNGLGNDPRARAVSVPLGGVRVAREFLQCRFLPAGPWLLF 359
 Db 301 GVECRFRGRGAMAWEGEPHRLNGLGNDPRARAVSVPLGGVRVAREFLQCRFLPAGPWLLF 360
 QY 360 SETSFISDVVNSGALGGTFPPAPWMPGPPPTNFSSLEPRGQOPVAKPKSSSTAIL 419
 Db 361 SETSFISDVVNSGALGGTFPPAPWMPGPPPTNFSSLEPRGQOPVAKPKSSSTAIL 416
 QY 420 IGCVAIILLLLIIALLMLRHLWRLLSKAERRVLEELTVHLSVPGDTILINRPGPR 479
 Db 417 IGCVAIILLLLIIALLMLRHLWRLLSKAERRVLEELTVHLSVPGDTILINRPGPR 476
 QY 480 EPPYQEPFRGRGPPHSPACVPGNSALLSNPAYRLLATYARPPRGPPPTPAWAKPTN 539
 Db 477 EPPYQEPFRGRGPPHSPACVPGNSALLSNPAYRLLATYARPPRGPPPTPAWAKPTN 536
 QY 540 TQAYSGDYMEPEKPGAPLLPPPPQNSVPHYAEADIVTLQGTGNTYAVPALPPGAVGDS 599
 Db 537 TQAYSGDYMEPEKPGAPLLPPPPQNSVPHYAEADIVTLQGTGNTYAVPALPPGAVGDS 596
 QY 600 PPRVDFPRSRRLRFKEKLGEGQGFGEVHLCVEVDSQDLVSLDFPLNVRKHGPHLLVAVKILRP 659

597 PPRVDFPRSRURFKKLGEGGEGVHUCEVEDPDQDLVTSDFPISVQKGGPLLVAVILRP 656
 660 DATNARNDFLKEVKIMSRUKDPNIRLLGVCVQDDPCLCMITDMYMGDLNQFLSAHOLE 719
 657 DATNARNDFLKEVKIMSRUKDLNIRLLGVCVQDDPCLCMITDMYMGDLNQFLSAHOLE 716
 720 DAAEAGPAGGQAAGPTISYPMLLHVAAGTASGRMYLATLNFVHRDLATRNCLVGENET 779
 717 NKVTOGLPGDRESQDPTISYPMLLHVAAGTASGRMYLATLNFVHRDLATRNCLVGENET 776
 780 IKIADFGMSRLNLYAGDYRVRQGRAVLPIRMWAMECILMGKFTTASDVWAFGVTLWEVLM 839
 777 IKIADFGMSRLNLYAGDYRVRQGRAVLPIRMWAMECILMGKFTTASDVWAFGVTLWEVLM 836
 840 CRAOPFGGLTDEQVIENAGAEFFRDQGRQVYLSRPPACPGQLYELMLRCWSRESEQRPPFS 899
 837 CRSOPFGGLTDEQVIENAGAEFFRDQGRQVYLSRPPACPGQLYELMLRCWSRESEQRPPFS 896
 900 QLRFLAEDALNTV 913
 897 QLRFLAEDALNTV 910
 RESULT 3
 EDDI_MOUSE STANDARD; PRT; 911 AA.
 AC Q03146;
 DT 01-OCT-1994 (REL. 30, CREATED)
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE EPITHELIAL DISCOIDIN DOMAIN RECEPTOR 1 PRECURSOR (EC 2.7.1.112)
 DE (TYROSINE-PROTEIN KINASE CAK) (CELL ADHESION KINASE) (TYROSINE KINASE
 DE DDR) (DISCOIDIN RECEPTOR TYROSINE KINASE) (PROTEIN-TYROSINE KINASE
 DE MPK-6).
 GN EDDR1 OR CAK OR MPK6.
 OS MUS MUSCULUS (MOUSE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.
 [1]
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
 RC STRAIN-C57BL;
 RA MEDLINE; 96204002.
 RA PEREZ J.L., JING S.O., WONG T.W.;
 RT "Identification of two isoforms of the Cak receptor kinase that are
 RT coexpressed in breast tumor cell lines."
 RL ONCOGENE 12:1469-1477(1996).
 RN [2]
 RP SEQUENCE OF 766-822 FROM N.A.
 RC STRAIN-C57BL; TISSUE=EMBRYONIC BRAIN;
 RX MEDLINE; 93096484.
 RA GILARDI-HEBENSTREIT P., NIETO M.A., FRAIN M., MATTEI M.-G.,
 RA CHESTIER A., WILKINSON D.G., CHARNAY P.;
 RT "An Eph-related receptor protein tyrosine kinase gene segmentally
 RT expressed in the developing mouse hindbrain."
 RL ONCOGENE 7:2499-2506(1992).
 CC - FUNCTION MAY BE INVOLVED IN CELL-CELL INTERACTIONS AND
 CC RECOGNITION (BY SIMILARITY).
 CC - CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
 CC PROTEIN TYROSINE PHOSPHATE.
 CC - SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC - TISSUE SPECIFICITY: THE PREDOMINANT ISOFORM CAK I, IS EXPRESSED IN
 CC DEVELOPING EMBRYO AND ADULT BRAIN; CAK II IS EXPRESSED IN VARIOUS
 CC EPITHELIAL CELLS.
 CC - ALTERNATIVE PRODUCTS: TWO ISOFORMS ARE PRODUCED BY ALTERNATIVE
 CC SPLICING. THE LONG FORM (CALLED CAK I AND SHOWN HERE) AND THE
 CC SHORT FORM (CALLED CAK II) ONLY DIFFER BY THE ABSENCE OF A 37
 CC RESIDUES SEGMENT.
 CC - SIMILARITY: WITH THE CONSERVED CATALYTIC DOMAINS OF TYR-PROTEIN
 CC KINASES.
 CC - SIMILARITY: CONTAINS 1 F5/8 TYPE C DOMAIN.
 CC - SIMILARITY: BELONGS TO THE INSULIN RECEPTOR FAMILY OF TYROSINE-
 CC PROTEIN KINASES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 DR EMBL; L57509; G1161063;
 DR EMBL; X57240; G53198;
 DR PIR; S30502; S30502;
 DR MGI; 99216; CAK.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.
 DR PROSITE; PS01285; FA58C_1; 1.
 DR PROSITE; PS01286; FA58C_2; 1.
 DR PFAM; PF00069; pkinase; 1.
 DR PFAM; PF00754; F5_F8_type_C; 1.
 DR HSP; P11362; 1FC1.
 KW TRANSFERASE; TYROSINE-PROTEIN KINASE; GLYCOPROTEIN; SIGNAL;
 KW PHOSPHORYLATION; TRANSMEMBRANE; RECEPTOR; ATP-BINDING;
 KW ALTERNATIVE SPLICING.
 FT SIGNAL 1 19
 FT CHAIN 20 911
 FT DOMAIN 20 414
 FT DOMAIN 415 441
 FT DOMAIN 442 911
 FT DOMAIN 32 186
 FT DOMAIN 379 413
 FT DOMAIN 474 599
 FT DOMAIN 608 903
 FT NP_BIND 614 622
 FT BINDING 653 653
 FT ACT_SITE 764 764
 FT DISULFID 32 186
 FT MOD_RES 511 511
 FT MOD_RES 790 790
 FT MOD_RES 794 794
 FT MOD_RES 795 795
 FT CARBOHYD 213 213
 FT CARBOHYD 262 262
 FT CARBOHYD 372 372
 FT CARBOHYD 392 392
 FT VARSPIC 503 539
 FT SEQUENCE 911 AA; 101160 MW; 3DCBB321 CRC32;
 SQ
 Query Match 93.2%; Score 4568; DB 1; Length 911;
 Best Local Similarity 93.3%; Pred. No. 8.3e-226;
 Matches 854; Conservative 20; Mismatches 35; Indels 6; Gaps 3;
 QY 1 MPEALSS-LLLLLVASGDADMKHFDPAKCRVALGMDRTIPDSDISASSSSDSTAA 59
 DB 1 MGTGLTSLLLLLLVLTIGDADMKHFDPAKCRVALGMDRTIPDSDISASSSSDSTAA 60
 QY 60 RHRSLSSDGDGAWCPAGVPKPEEYLQVDLRLHLVALVGTQGRHAGGLGKFSRSYR 119
 DB 61 RHRSLSSDGDGAWCPAGVPKPEEYLQVDLRLHLVALVGTQGRHAGGLGKFSRSYR 120
 QY 120 LRYSDRGWRMCKRWGQEVISGNEDEPVGVLKDLGPPMVARLVRYPRADRVMSVCLR 179
 DB 121 LRYSDRGWRMCKRWGQEVISGNEDEPVGVLKDLGPPMVARLVRYPRADRVMSVCLR 180
 QY 180 VELYGLNRDGLLSVTAPVQGMYSLEA-VYLNDSYDGHVVGGLQGLQLADGVVGL 238
 DB 181 VELYGLNRDGLLSVTAPVQGMYSLEA-VYLNDSYDGHVVGGLQGLQLADGVVGL 240
 QY 239 DDFRKSQELRWPGDYDYGWSNHSFSSGVVEMEFDFRLRAFOAMQVHCNNHHTLGARLP 298

```
Db 241 DFRQSOELRVWPGYDYVGSWNSQSFPTGVMEFEFDRLRTFTQWVHCNNHMTLGARLP 300
QY 299 GGYECFRGRGPMAMGECEPMRHNGLGNLGDPRARAVSVPLGGRVAFRLQCRFLFAGPWLL 358
Db 301 GGYECFRGRGPMAMGECEPMRHNGLGNLGDPRARAVSVPLGGRVAFRLQCRFLFAGPWLL 360
QY 359 FSEISFISDVVNNSSPALGGTTPPAPWPPGPPPTNFSSLELEPRGQOQVAKPEGSPAT 418
Db 361 FSEISFISDVVNNSSPALGGTTPPAPWPPGPPPTNFSSLELEPRGQOQVAKPEGSPAT 416
QY 419 LIGCLVAIILLLLLLIALMLRHLWRLLSKAERRVLEELTVHLSVPGDTILINRGP 478
Db 417 LIGCLVAIILLLLLLIALMLRHLWRLLSKAERRVLEELTVHLSVPGDTILINRGP 476
QY 479 REPPPYQEPGRGPMAMGECEPMRHNGLGNLGDPRARAVSVPLGGRVAFRLQCRFLFAGPWLL 538
Db 477 REPPPYQEPGRGPMAMGECEPMRHNGLGNLGDPRARAVSVPLGGRVAFRLQCRFLFAGPWLL 536
QY 539 NTQAYSGDYMEPEKPGAPLLPPPPQNSVPHYADIVTLQGVGTGNTYAVPALPGAVGD 598
Db 537 NTQAYSGDYMEPEKPGAPLLPPPPQNSVPHYADIVTLQGVGTGNTYAVPALPGAVGD 596
QY 599 GPRVDFPRSRLEKEKLGEGQGEVHLCVDSQDLVSLDPLNVRKGHPLLVAKILR 658
Db 597 GPRVDFPRSRLEKEKLGEGQGEVHLCVDSQDLVSLDPLNVRKGHPLLVAKILR 656
QY 659 PDATKARNDFLEKVMISRLKDPNIIRLIGVCVQDDPLCMITDYMENGLNOFSAHOL 718
Db 657 PDATKARNDFLEKVMISRLKDPNIIRLIGVCVQDDPLCMITDYMENGLNOFSAHOL 716
QY 719 EDKAAEGDGGQAAOGPTTISYPMLLHVAQAIAAGMRYLATLNFVHRDLATRNCLVGENF 778
Db 717 ENKATQGLSGDTSQDPTISYPMLLHVAQAIAAGMRYLATLNFVHRDLATRNCLVGENF 776
QY 779 TIKIADFGMSRLNLYAGDYRVQRAVLPIRMAWECILMGKFTTASDVWAFVTLWEVLM 838
Db 777 TIKIADFGMSRLNLYAGDYRVQRAVLPIRMAWECILMGKFTTASDVWAFVTLWEVLM 836
QY 839 LCRAQPGQLTDEQVIENAGEFFRDGQVYLSRPPACPOGLYELMRCWSREORPPF 898
Db 837 LCRAQPGQLTDEQVIENAGEFFRDGQVYLSRPPACPOGLYELMRCWSREORPPF 896
QY 899 SOLHRLFAEDALNTV 913
Db 897 AQLHRLFAEDALNTV 911

RESULT 4
ID TRK3_HUMAN STANDARD; PRT; 855 AA.
AC Q16832;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DE RECEPTOR PROTEIN-TYROSINE KINASE TKT PRECURSOR (EC 2.7.1.112)
DE (TYROSINE-PROTEIN KINASE TYRO 10) (NEUROTROPHIC TYROSINE KINASE,
DE RECEPTOR-RELATED 3).
GN NRK3 OR TKT OR TYRO10.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=HEART, AND THYMUS;
RA MEDLINE; 94067796.
RA KARN T., HOLTRICH U., BRAEUNINGER A., BOEHME B., WOLF G.,
RA RUEBSAMEN-WAIGMANN H., STREBHARDT K.;
RT "Structure, expression and chromosomal mapping of tkt from man and
RT mouse: a new subclass of receptor tyrosine kinases with a factor
RT vlii-like domain";
RL ONCOGENE 8:3433-3440(1993).
CC -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE =
CC ADP + PROTEIN TYROSINE PHOSPHATE.
```

```
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- TISSUE SPECIFICITY: THE MAJOR 10 KD TRANSCRIPT IS EXPRESSED IN
CC HIGH LEVELS IN HEART AND LUNG, LESS IN BRAIN, PLACENTA, LIVER,
CC SKELETAL MUSCLE, PANCREAS, AND KIDNEY.
CC -!- ALTERNATIVE PRODUCTS: DIFFERENT TRANSCRIPTS ARE DERIVED FROM ONE
CC GENE.
CC -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 F5/8 TYPE C DOMAIN.
CC -!- SIMILARITY: BELONGS TO THE INSULIN RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC MIM: 191311;
CC EMBL: X74764; G433338;
CC
CC PROSITE; PS00109; PROTEIN KINASE_TYR_1;
CC PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1;
CC PROSITE; PS01285; FA58C_1; 1;
CC PROSITE; PS01286; FA58C_2; 1;
CC PROSITE; PS50011; PROTEIN KINASE_DOM; 1;
CC PFAM; PF00069; pkinase; 1;
CC PFAM; PF00754; F5_F8_type_C; 1;
CC HSP; P11362; IFGI.
CC TRANSFERASE; TYROSINE-PROTEIN KINASE; GLYCOPROTEIN; SIGNAL;
CC PHOSPHORYLATION; TRANSMEMBRANE; RECEPTOR; ATP-BINDING.
CC SIGNAL 1 21
CC CHAIN 22 855
CC RECEPTOR PROTEIN-TYROSINE KINASE TKT.
CC FT DOMAIN 22 399
CC EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 400 421
CC POTENTIAL.
CC FT DOMAIN 422 855
CC CYTOPLASMIC (POTENTIAL).
CC FT DOMAIN 30 185
CC F5/8 TYPE C.
CC FT DOMAIN 563 7
CC PROTEIN KINASE.
CC FT NP_BIND 569 577
CC ATP (BY SIMILARITY).
CC FT BINDING 608 608
CC ATP (BY SIMILARITY).
CC FT ACT_SITE 710 710
CC BY SIMILARITY.
CC FT DISULFID 30 185
CC BY SIMILARITY.
CC FT MOD_RES 740 740
CC PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
CC FT CARBOHYD 121 121
CC POTENTIAL.
CC FT CARBOHYD 213 213
CC POTENTIAL.
CC FT CARBOHYD 261 261
CC POTENTIAL.
CC FT CARBOHYD 280 280
CC POTENTIAL.
CC FT CARBOHYD 372 372
CC POTENTIAL.
CC SQ SEQUENCE 855 AA; 96752 MW; 2623D841 CRC32;

Query Match 49.38; Score 2415; DB 1; Length 855;
Best Local Similarity 52.1%; Pred. No. 2.7e-116;
Matches 482; Conservative 118; Mismatches 227; Indels 98; Gaps 15;

QY 3 PEALSSLLLLLVASGDADMGKGFDPKACRYALGMDRTIPDSISASSWSDDTAARHS 62
Db 5 PMLLVFLLLPILS---SARAQVNPALCRYPGLMSGQIPEDITASSQWSESTAAYG 61
QY 63 RLESDGDGAWCPAGSVFPKE--EYLVQDLQLRHLVALVGTQGRHAGGLGKEFSRSLR 121
Db 62 RLDSEGDGAWCPPEIPVEPDDILKEFLQIDLTLHTLFTLVGTQGRHAGGLGKEFSR 121
QY 122 YSRDGRRWGKDRWGQEVISGNDPEGVVLDKLGPPMVARLVRFYPRADRVMSVCLRVE 181
Db 122 YSRDGRRWGKDRWGQEVISGNDPEGVVLDKLGPPMVARLVRFYPRADRVMSVCLRVE 181
QY 182 LYGLCWRGLLSYTPAVGQTMVL--SEAVYLNDSYDCHTGGGLQYGGGLADGVVGLD 239
Db 182 LYGLCWRGLLSYTPAVGQTMVL--SEAVYLNDSYDCHTGGGLQYGGGLADGVVGLD 239
QY 182 LYGLCWRGLLSYTPAVGQTMVL--SEAVYLNDSYDCHTGGGLQYGGGLADGVVGLD 240
Db 182 LYGLCWRGLLSYTPAVGQTMVL--SEAVYLNDSYDCHTGGGLQYGGGLADGVVGLD 240
QY 240 DFRKSOELRVWPGYDYVGSWNSHFSFGYVMEFEFDRLRAFAQMVHCNNHMTLGARLP 299
```

```
Db 241 DTQTHVHWPCYDVGWRNESATNGIEMFEEDRIKRNFTTKVHCNNMFAGVKIEK 300
Qy 300 GVECFRRGPGAMWEGEPHNRHNLGNLGDPRARAVSPGLGRVARFLQCRFLFAGPWLLF 359
Db 301 EVQCYF-RSEASEWEPNALSPLVLDVNPASRVTVPLHHRMASAIKCOVHFADTWMF 359
Qy 360 SEISFISD-VVNSSPALGGTFPPAPWPPGPPPTNFSSLEPRGQOPVAKPGSPAI 418
Db 360 SEITFQSDAAMYNSEAL-----PTSP-----MAPTTVDPMKLVDSNTRI 400
Qy 419 LGICLVAILLLLLIALLMLRLHWRLLSKAERVLEEELTVHLSVPGDTILNNR-P 476
Db 401 LGICLVAILFILLIIVILWRQFQWQKLEKASRMDDDEMTVSLSPSSFMFNNRNS 460
Qy 477 GPREP-----PPYQEP-PRGNPPHSPACVPNGSALLSNPVRLLILATYARP 523
Db 461 SPSEQGSNTYDRIFFLPDPQEP-----SRLRKLPF----- 495
Qy 524 PRGPGPTPAWAKPTNTQAYSGDYWEPEKPGAPLLPPPPQNSVPHYAEADIVTLQVGTG 583
Db 495 -----APGEESGCGSVVQVPQSGP-----EGVPHYAEADIVNLQVGTG 535
Qy 584 NTYAVPALPPGAVGDPGRV-DEPRSRLRFXEKLGEQGFGEVHLCEVDSQDLVSLDFPL 642
Db 536 NTYSPVATMDLLSGKVAVAEFFPKLLTFKELGEGQGFGEVHLCEVEGMEKFKDKDFAL 595
Qy 643 NVRGHPLLVAVKILRPDATKNARNDFLKEVKIMSLRKLDPNIIIRLLGVCVQDDPLCMITD 702
Db 596 DVSAQNPVLVAVKMLRADANKNARNDFLKEIKIMSLRKLDPNIIIRLLGVCITDPLCMITE 655
Qy 703 YWENGDLNQFLSAHQLEDAEAGAPGCGQAAGQGTISYPMLLHVAAGIAGMRYLATLNF 762
Db 656 YWENGDLNQFLSRHE-----PPNSSSDVTVSYTNLKFMTQIAGSMKYLSSLNF 706
Qy 763 VHRDLATNCLVGNFTIKIADFGMSRLYAGDYRVQGRAVLPPIRWNAECILMGKFTT 822
Db 707 VHRDLATNCLVGNFTIKIADFGMSRLYAGDYRVQGRAVLPPIRWMSIESILGKFTT 766
Qy 823 ASDVWAFGVTLMWELMCLRAQPGFOLTDEQVIENAGFEFFRDQGRQVYLSRPPACPGLYE 882
Db 767 ASDVWAFGVTLMWETFTFCQEPYQSLSDEQVIENAGFEFFRDQGRQVYLSRPPACPGLYE 826
Qy 883 LMLKCSWRESQRPFPFSLHFLAE 907
Db 827 LMLKSWRDRDTKRNPSFOEIHLLQ 851

RESULT 5
ID TRK3_MOUSE STANDARD; PRT; 854 AA.
AC Q62371;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DE RECEPTOR PROTEIN-TYROSINE KINASE TKT PRECURSOR (EC 2.7.1.112)
DE (TYROSINE-PROTEIN KINASE TYRO 10) (NEUROTROPHIC TYROSINE KINASE,
DE RECEPTOR-RELATED 3).
CN TRK3 OR TKT OR TYRO10.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 94067796.
RA KARN T., HOLTRICH U., BRAEUNINGER A., BOEHME B., WOLF G.,
RA RUEBSAMEN-WAIGMANN H., STREBHARDT K.;
RT "Structure, expression and chromosomal mapping of TKT from man and
RT mouse: a new subclass of receptor tyrosine kinases with a factor
RL viii-like domain."
RL ONCOGENE 8:3433-3440(1993).
RN [2]
RP SEQUENCE FROM N.A.
```

```
RC STRAIN-C57BL/6; TISSUE=BRAIN;
RX MEDLINE; 94151011.
RA LAI C., LEMKE G.E.;
RT "Structure and expression of the Tyro 10 receptor tyrosine kinase.";
RL ONCOGENE 9:877-883(1994).
CC -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE -
CC ADP + PROTEIN TYROSINE PHOSPHATE.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- TISSUE SPECIFICITY: WIDELY EXPRESSED; HIGH LEVELS IN SKELETAL
CC MUSCLE, HEART, CNS, AND KIDNEY; LESS IN OTHER TISSUES. THE MAJOR
CC 10 KD TRANSCRIPT IS EXPRESSED IN HIGH LEVELS IN HEART AND LUNG,
CC LESS IN BRAIN AND TESTIS.
CC -!- ALTERNATIVE PRODUCTS: DIFFERENT TRANSCRIPTS ARE DERIVED FROM ONE
CC GENE.
CC -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 F5/8 TYPE C DOMAIN.
CC -!- SIMILARITY: BELONGS TO THE INSULIN RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X76505; G435162; ALT_INIT.
CC MGD; MGI:97385; NTRK3.
CC PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
CC PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.
CC PROSITE; PS01285; FA58C.1; 1.
CC PROSITE; PS01286; FA58C.2; 1.
CC PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
CC PFAM; PF00089; Pkinase; 1.
CC PFAM; PF00754; F5_F8_type_C; 1.
CC HSP; P11362; IFGI.
CC -----
CC TRANSFERASE; TYROSINE-PROTEIN KINASE; GLYCOPROTEIN; SIGNAL;
CC PHOSPHORYLATION; TRANSMEMBRANE; RECEPTOR; ATP-BINDING.
CC -----
CC SIGNAL 1 21
CC CHAIN 22 854
CC DOMAIN 22 399
CC TRANSMEM 400 421
CC DOMAIN 422 854
CC DOMAIN 30 185
CC DOMAIN 563 577
CC NP_BIND 569 577
CC BINDING 608 608
CC ACT_SITE 709 709
CC DISULFID 30 185
CC MOD_RES 739 739
CC CARBOHYD 121 121
CC CARBOHYD 213 213
CC CARBOHYD 261 261
CC CARBOHYD 280 280
CC CARBOHYD 372 372
CC SEQUENCE 854 AA; 96482 MW; B78B6551 CRC32;

Query Match 49.2%; Score 2413; DB 1; Length 854;
Best Local Similarity 52.3%; Pred. No. 3.4e-116;
Matches 481; Conservative 119; Mismatches 220; Indels 100; Gaps 15;

Qy 9 LLLLLLVASGDADMKGHEDPAKRYALGMQDRTTPDSDISASSWSSTAAHRSLSSD 68
Db 10 VLLLLLLILGSA--KAQVNPACRYPLGMSGGHIPDEDITASSQWSESTAAKYLRLDSE 67
Qy 69 GDGAWCPAGSVFPKE-EYELQVDLQRLHLVALVGTQGRHAGLGKFKFSRSLRYLRSYDR 127
Db 68 GDGAWCPPIPVQDDLKKEFLQIDLRLTHFLVGTQGRHAGGIEFAPMKYKINSRDS 127
Qy 128 RWMGWKDRWGQEVISGNEDEPVGVLKDLGPPMVARLVRFYPRADRVMSVCLRLVELYGLW 187
```

```

128 RWISWRHCKQVLDNSNPYDFLKDLEPPVAVRVLIPVTDHSMNVCMLVGLGCVW 187
188 RGGLSYTAVGQTMYL--SEAVLYNDSTYDGTGVLGQYGLQGLADGVGLDFRKSQ 245
188 LDGLVSNAPAGQOQVLPGLGSIYIYNDSDYDG-AGVYSMTGSLQGLTGVSGLDFTQTH 246
246 ELRWPGYDVGKSHNSFSGVYEMEFEDRLRAFOAMQVHNNMHTLGLARLPGGVCRF 305
247 EYHWPGYDVGWRNESATNGFIEINFEDRIETNMKVHCNNMFAGVKIFKEVQCYF 306
306 RRGPMAMWEGEPMRHLGNLGDPRARAVSVPLGGVAREFLOCFLEAGPMLFSEIFL 365
307 -RSEASEWPTAVYFLVLDVNPVSARFVTVPLHHRMSAICQHFADTWNMFSEITFQ 365
366 SD--VYNNSPALGGTTPAPMWPDPPTNFSSLELEPRGQOPVAKPEGSTAILIGCL 423
366 SDAAMYNNS--GALPTSP-----MAPTTDPMKLVKDDSNTRILIGCL 405
424 VAILLLLLIALLMRLHWRLLSKAERVLEELTVHLSVPGDTILINR-----PQPR 479
406 VALIFILLIIVILWQFQWMLKASRRMLDDEMTVSLSPSESMFNNRRSSPSEQ 465
480 EP-----PPYQEPFRGNPHSAPCVPGNSALLSNPAYRLLATVAPRPGG 528
466 ESNSTYDRIFPLRDPQEP-----SLRLKLPF----- 495
529 PPTPAWAKTNTQAVSGDYMEPEKPGAPLPPPPQNSVPHYAEADIVTLQGTGNTYAV 588
495 -----APGEESGCGVVKPAQNGP-----EGVPHYAEADIVNLQGTGNTYCV 540
589 PALPGAVGDGPRV--DFPRSLRKEKLGEGQFGEVHLCVDSQDVLVSLDFPLNVRKG 647
541 PATMDLLSGKDVAVEEFPKLLAKEKLGEGQFGEVHLCVDSQDVLVSLDFPLNVRKG 600
648 HPLLVAKILRPAVKNAERDFELKEVKIMSLKDPNIIRLLGVCVQDDPLCMITDYMENG 707
601 QPVLVAVKMLRADANKNAERDFELKEVKIMSLKDPNIIRLLAVCTEDPLCMITDYMENG 660
708 DLNQFLSAHQLEKAEAGAPGQAAQGTTSYPMLLHVAQIASGRMYLATLNFVHRDL 767
661 DLNQFLSRHEPLSSGSSDA-----TVSYANKFMATQIASGRMYLATLNFVHRDL 710
768 ATRNCLVGENFTIKTADFGMSRLNLAGDYRYVQGRAVPIRMAWECILMGKFTASDVW 827
711 ATRNCLVGNKTYIKTADFGMSRLNLAGDYRYVQGRAVPIRMAWESILLGKFTASDVW 770
828 AFGVTLWEVLMCLRAQPFQGLTDEQVIENAGFEFFRQDQROVYLSRPPACPGLYELMLRC 887
771 AFGVTLWEVTFECQEQPYSQLSDEQVIENAGFEFFRQDQROVYLSRPPACPGLYELMLRC 830
888 WSRESEORPPFSQHLRFIAE 907
831 WRRETKHRPSFOEIHLLIQ 850

RESULT 6
TRKB_CHICK
AC Q91010; STANDARD; PRT; 818 AA.
DT 15-JUL-1998 (REL. 36, CREATED)
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE BDNF / NT-3 GROWTH FACTORS RECEPTOR PRECURSOR (EC 2.7.1.112) (TRKB
DE TYROSINE KINASE) (TRKB-B).
OS GALLUS GALLUS (CHICKEN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ARCHOSAURIA; AVES;
OC NEOGNATHAE; GALLIFORMES; PHASIANIDAE; PHASIANINAE; GALLUS.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE; 95047511.

```

```

RA VINH N., ERDMANN K., HEUMANN R.;
RT "Cloning and sequence analysis of a cDNA encoding a novel truncated
RT form of the chicken TrkB receptor.";
RL GENE 149:383-384(1994).
RP SEQUENCE FROM N.A.
RX MEDLINE; 94116452.
RA DECHANT G., BIFFO S., OKAZAWA H., KOLBECK R., POTTGISSER J.,
RA BARDE Y.A.;
RT "Expression and binding characteristics of the BDNF receptor chick
RT trkB";
RL DEVELOPMENT 119:545-558(1993).
CC -1- FUNCTION: RECEPTOR FOR BRAIN-DERIVED NEUROTROPHIC FACTOR (BDNF),
CC NEUROTROPHIN-3 AND NEUROTROPHIN-4/5 BUT NOT NERVE GROWTH FACTOR
CC (NGF). INVOLVED IN THE DEVELOPMENT AND/OR MAINTENANCE OF THE
CC NERVOUS SYSTEM. THIS IS A TYROSINE-PROTEIN KINASE RECEPTOR. KNOWN
CC SUBSTRATES FOR THE TRK RECEPTORS ARE SHC, PI-3 KINASE, AND PLC-
CC GAMMA-1 (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC PROTEIN TYROSINE PHOSPHATE.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- PTM: LIGAND-MEDIATED AUTO-PHOSPHORYLATION.
CC -1- ALTERNATIVE PRODUCTS: A TRUNCATED NON-CATALYTIC FORM WHICH
CC FUNCTION REMAINS STILL UNCLEAR IS PRODUCED BY ALTERNATIVE
CC SPLICING.
CC -1- SUBUNIT: EXISTS IN A DYNAMIC EQUILIBRIUM BETWEEN MONOMERIC (LOW
CC AFFINITY) AND DIMERIC (HIGH AFFINITY) STRUCTURES (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE INSULIN RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES. HIGHEST SIMILARITY WITH OTHER TRK-TYPE RECEPTORS.
CC -1- SIMILARITY: THE REPEATED LEUCINE-RICH (LRR) SEGMENT IS FOUND IN
CC MANY PROTEINS. NUMBER IN THIS PROTEIN: 2.
CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CONTAINS
CC TWO C2-LIKE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; X77251; G472934; -
EMBL; X77252; G472936; -
EMBL; X74109; G407799; -
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.
DR PFAM; PF00047; ig; 1.
DR PFAM; PF00069; pkinase; 1.
DR PFAM; PF00560; LRR; 1.
DR HSP; P11362; IFGI.
KW TRANSFERASE; TYROSINE-PROTEIN KINASE; TRANSMEMBRANE; ATP-BINDING;
KW PHOSPHORYLATION; RECEPTOR; GLYCOPROTEIN; NEUROGENESIS; SIGNAL;
KW LEUCINE-REPEAT; REPEAT; ALTERNATIVE SPLICING; IMMUNOGLOBULIN FOLD.
FT SIGNAL 1 31
FT CHAIN 32 818
FT DOMAIN 32 426
FT TRANSMEM 427 450
FT DOMAIN 451 818
FT REPEAT 71 116
FT REPEAT 95 116
FT DOMAIN 213 269
FT DOMAIN 300 364
FT NP_BIND 540 548
FT BINDING 568 568
FT ACT_SITE 672 672
FT MOD_RES 512 512
FT MOD_RES 698 698
FT MOD_RES 702 702
FT CHAIN 32 818
FT DOMAIN 32 426
FT TRANSMEM 427 450
FT DOMAIN 451 818
FT REPEAT 71 116
FT REPEAT 95 116
FT DOMAIN 213 269
FT DOMAIN 300 364
FT NP_BIND 540 548
FT BINDING 568 568
FT ACT_SITE 672 672
FT MOD_RES 512 512
FT MOD_RES 698 698
FT MOD_RES 702 702

```


476 SASPLHHISNGSNTSPSSSEGGPDAVIIGMTKIPVIEPNQYFQITNSQLKPDFTFVQHIKRR 533S

QY 609 RLRFKEKLGEGQFEVHLCE---VDSQDQLVSLDFPLNVRKRGHPLVAVKILRPDATKNA 665S

Db 536 NIVLKRLEGAGFKVLAECYNLCPEQD-----KILVAVKTLK-DASDNA 580S

QY 666 RNDELKEVKIMSLKDPNIIRLLGVGVQDDPLCHITDYMENGOLNOFLSAHQLED-KAAE 724S

Db 591 RKDFHRAELLTNLQHEHIVKFGVGVCEGDPILMVFEYMKHGDLNKLFLRAHGDPDAVLMAE 640S

QY 725 GAPDGGQAAQGP-T-ISYPMLLHVAQAQIASGRMYLATLNFVHRDLATRNCLVGENETIKTA 763S

Db 641 GNP-----PTELTSQMLHQAQIAAGVYLAESHOFVHRDLATRNCLVGENLLVKIG 692S

QY 784 DFGMSRLNLYAGDYRVROGRAVLPIRNMWAEICILMGKFTTASDYVAFGYTLWEVLMLCRAQ 843S

Db 593 DFGMSRDVYSTDYRVGVGHGTMPIRWMPPEISIMYRKFTEDESQVSLGVVLWEFTYVK-Q 751S

QY 844 PFGQLTDEQVIENAGAEFFRQGVYLSRRPPACPGIYELMLRCWSESEORPPFSQHLR 903S

Db 752 PWYQLSNNEVIECI-----TQGR-VLQRPRTCPQEVIELMGLWCQOREPHTRKNIKSIHT 804S

QY 904 FL 905

Db 805 LL 806

RESULT 8

ID	TRKB_RAT	STANDARD;	PRT;	821 AA.
AC	Q63604;	Q63605;	Q63606;	
DT	01-NOV-1997	(REL. 35, CREATED)		
DT	01-NOV-1997	(REL. 35, LAST SEQUENCE UPDATE)		
DT	15-JUL-1998	(REL. 36, LAST ANNOTATION UPDATE)		
DE	BNF / NT-3	GROWTH FACTORS RECEPTOR PRECURSOR (EC 2.7.1.112) (TRKB		
DE	TYROSINE KINASE)	(GP145-TRKB / GP95-TRKB) (TRK-B).		
GN	NRK2 OR TRKB.			
OS	RATTUS NORVEGICUS (RAT).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;			
OC	RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.			
[1]				
RP	SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.			
RC	TISSUE-CEREBELLUM;			
RC	MEDLINE; 91094826.			
RA	MIDDLEMAS D.S., LINDBERG R.A., HUNTER T.;			
RT	"trkb, a neural receptor protein-tyrosine kinase: evidence for a			
RT	full-length and two truncated receptors.";			
RL	MOL. CELL. BIOL. 11:143-153(1991).			
RL	[2]			
RP	PHOSPHORYLATION SITES.			
RP	MEDLINE; 94149017.			
RA	MIDDLEMAS D.S., MEISENHOLDER J., HUNTER T.;			
RT	"Identification of Trkb autophosphorylation sites and evidence that			
RT	phospholipase C-gamma 1 is a substrate of the Trkb receptor.";			
RL	J. BIOL. CHEM. 269:5458-5466(1994).			
CC	-!- FUNCTION: RECEPTOR FOR BRAIN-DERIVED NEUROTROPHIC FACTOR (BDNF),			
CC	NEUROTROPHIN-3 AND NEUROTROPHIN-4/5 BUT NOT NERVE GROWTH FACTOR			
CC	(NGF). INVOLVED IN THE DEVELOPMENT AND/OR MAINTENANCE OF THE			
CC	NERVOUS SYSTEM. THIS IS A TYROSINE-PROTEIN KINASE RECEPTOR. KNOWN			
CC	SUBSTRATES FOR THE TRK RECEPTORS ARE SHC, PI-3 KINASE, AND PLC-			
CC	GAMMA-1.			
CC	-!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +			
CC	PROTEIN TYROSINE PHOSPHATE.			
CC	-!- SUBUNIT: EXISTS IN A DYNAMIC EQUILIBRIUM BETWEEN MONOMERIC (LOW			
CC	AFFINITY) AND DIMERIC (HIGH AFFINITY) STRUCTURES (BY SIMILARITY).			
CC	-!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.			
CC	-!- TISSUE SPECIFICITY: THE DIFFERENT FORMS ARE DIFFERENTIALLY			
CC	EXPRESSED IN VARIOUS CELL TYPES. THE T2 ISOFORM IS PRIMARILY			
CC	EXPRESSED IN NEURONS.			
CC	-!- ALTERNATIVE PRODUCTS: AT LEAST THREE FORMS ARE PRODUCED BY			
CC	ALTERNATIVE SPLICING; THESE ARE: GP145-TRKB, T1 (GP95-TRKB) AND			
CC	T2.			
CC	-!- PTM: LIGAND-MEDATED AUTO-PHOSPHORYLATION.			

-!- SIMILARITY: BELONGS TO THE INSULIN RECEPTOR FAMILY OF TYROSINE-
PROTEIN KINASES. HIGHEST SIMILARITY WITH OTHER TRK-TYPE RECEPTORS.
-!- SIMILARITY: THE REPEATED LEUCINE-RICH (LRR) SEGMENT IS FOUND IN
MANY PROTEINS. NUMBER IN THIS PROTEIN: 2.
-!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CONTAINS
TWO C2-LIKE DOMAIN.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; M55291; G207474; -;
EMBL; M55292; G207476; -;
EMBL; M55293; G207478; -;
PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.
PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
PFAM; PF00047; 1g; 1.
PFAM; PF00069; pkinase; 1.
PFAM; PF00560; LRR; 1.
HSP; P11362; IFGI.
TRANSFERASE; TYROSINE-PROTEIN KINASE; TRANSMEMBRANE; ATP-BINDING;
KW PHOSPHORYLATION; RECEPTOR; GLYCOPROTEIN; NEUROGENESIS; SIGNAL;
KW LEUCINE-REPEAT; REPEAT; IMMUNOGLOBULIN FOLD; ALTERNATIVE SPLICING.
FT SIGNAL 1 31
FT CHAIN 32 821
FT DOMAIN 32 429
FT TRANSMEM 430 453
FT DOMAIN 454 821
FT DOMAIN 72 93
FT REPEAT 96 117
FT REPEAT 117 117
FT DOMAIN 214 270
FT DOMAIN 301 365
FT DOMAIN 537 806
FT NP_BIND 543 551
FT BINDING 571 571
FT ACT_SITE 675 675
FT MOD_RES 515 515
FT MOD_RES 701 701
FT MOD_RES 705 705
FT MOD_RES 706 706
FT MOD_RES 816 816
FT SITE 515 515
FT SITE 816 816
FT SITE 816 816
FT CARBOHYD 57 67
FT CARBOHYD 95 95
FT CARBOHYD 121 121
FT CARBOHYD 178 178
FT CARBOHYD 205 205
FT CARBOHYD 241 241
FT CARBOHYD 254 254
FT CARBOHYD 280 280
FT CARBOHYD 325 325
FT CARBOHYD 338 338
FT CARBOHYD 411 411
FT VARSPPLIC 466 466
FT VARSPPLIC 477 821
FT VARSPPLIC 466 474
FT VARSPPLIC 475 821
SO SEQUENCE 821 AA; 92186 MW; 884222D2 CRC32;
Query Match 13.5%; Score 659.5; DB 1; Length 821;
Best Local Similarity 30.4%; Pred. No. 5.5e-27;

Matches 201; Conservative 72; Mismatches 188; Indels 201; Gaps 22;
QY 373 SPALGGTTPPPAPWPPGPPPTNFSSLEPRGQOPVAKPEGSPAILGICLVAILLLL 432
Db 217 SCVGGDPLPLYWDVGNLVSKHNEISHTOGSLRIINISSDDSGKQISCV----- 268
QY 433 IIALMLRLHRRLLSKAERVLBEELTVHLSV---PGDTILINRRPGPREPP----- 483
Db 268 -----AENLVGEDQDSVNLTVHFAPTIFEL-----ESPTSDHHC 302
QY 483 -PYQPRPRGNPPHAPCVNPGS-----ALLSNPA-----Y 513
Db 303 IPF---TVRGNPKPALQFYNGAILNESKYICTKIHTVNTHTYHGCLODNTTHNNGDY 359
QY 514 RLL-----ATYARPPRGPGPTPAWAKPTNTQAYSGDYMEP--E 551
Db 360 TLMKNEYKDERQISAHFMGRPGVDYETNPYPVLYEDWTPTDI-----GDITNSNE 415
QY 552 KPGAPLLPPPPONSYPHYAEDIVTLQV----- 581
Db 416 IPSTDVADQTNREHLSVYAVVYASVVGFCLLVMLLLKLARHSKFGMKGPASVISNDDD 475
QY 581 -----TGGNTYAVPALPGAVDGP---PRVDEP-----RS 608
Db 476 SASPLHHISNGSNTSPSSSEGGPDVILGKIPVIENQYFGITNSQLKPDFTVQHIKRH 535
QY 609 RLRFKEKLEGQGEVHLCE---VDSQDVLSDPLNVRKGHPLLVAVKILRPDATKNA 665
Db 536 NIVLRKELGEGAFGVFLAECYNLCPEQ-----KILVAVKTLK-DASDNA 580
QY 666 RNDFLKEVKIMRLKDPNIIRLLGVCVQDDPLCMITDYMENDLNQFLSAHQLED-KAAE 724
Db 581 RKDFHREAEELLNLQHEHIVKFYGVCEGDPIMVFEYMKHGDNLKFLRAHGPDAVLMAE 640
QY 725 GAPGDAQAGPT-ISYPMLLHVAQIASGRYLAFLNFVHRDLATRNCLVGENETIKIA 783
Db 641 GNP-----PIELTQSNLHIAQIAAGVYLASQHFVHRDLATRNCLVGENLLVIG 692
QY 784 DFGMSRLYAGDYRYVQGRAVLPIRMWAEICLMGKFTASDVAFGVTWLVLMCLCRAQ 843
Db 693 DFGMSRDVYSTDYRYVGGHTMLPIRMWPESTMYRKFTTESDWSLVGVVWEIFTYGR-Q 751
QY 844 PFQLTDEQVIENAGEFFRDQGRQVYLSRPPACPOGLVELMLRCWSRESEQPPFSQHR 903
Db 752 PWQLSNNEVIECI-----TOGR--VLRPRTCPQEVYELMLGWCQREPHTRKNKNIHT 804
QY 904 FL 905
Db 805 LL 806
RESULT 9
TRKB_HUMAN
ID TRKB_HUMAN STANDARD; PRT; 822 AA.
AC Q16620; Q16675;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE BDNF / NT-3 GROWTH FACTORS RECEPTOR PRECURSOR (EC 2.7.1.112) (TRKB
DE TYROSINE KINASE) (GP145-TRKB) (TRK-B).
GN NTRK2 OR TRKB
OS HOMO SAPIENS (HUMAN)
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A. (LONG FORM).
RC TISSUE-HIPPOCAMPUS;
RX MEDLINE; 9530922.
RA NAKAGAWARA A., LIU X.-G., IKEGAKI N., WHITE P.S., YAMASHIRO D.J.,
RA NYCUM L.M., BIEGEL J.A., BRODEUR G.M.;
RT "Cloning and chromosomal localization of the human TRK-B tyrosine
RT kinase receptor gene (NTRK2).";
RL GENOMICS 25:538-546(1995).

[2] SEQUENCE FROM N.A., AND PARTIAL SEQUENCE (LONG AND TRUNCATED FORMS).
RP TISSUE-BRAIN;
RC MEDLINE; 95123473.
RX SHELTON D.L., SUTHERLAND J., GRIPP J., CAMERATO T., ARMANINI M.P.,
RA PHILLIPS H.S., CARROLL K., SPENCER S.D., LEVINSON A.D.;
FT "Human trks: molecular cloning, tissue distribution, and expression
RT of extracellular domain immunoadhesins.";
RL J. NEUROSCI. 15:477-491(1995).
[3]
RN SEQUENCE FROM N.A. (TRUNCATED FORMS).
RP TISSUE=HIPPOCAMPUS;
RC MEDLINE; 95022162.
RX ALLEN S.J., DAWBARN D., ECKFORD S.D., WILCOCK G.K., ASHCROFT M.,
RA COLEBROOK S.M., FEENEY R., MACGOWAN S.H.;
FT "Cloning of a non-catalytic form of human trkB and distribution of
RT messenger RNA for trkB in human brain.";
RL NEUROSCIENCE 60:825-834(1994).
CC -1- FUNCTION: RECEPTOR FOR BRAIN-DERIVED NEUROTROPHIC FACTOR (BDNF),
CC NEUROTROPIN-3 AND NEUROTROPIN-4/5 BUT NOT NERVE GROWTH FACTOR
CC (NGF). INVOLVED IN THE DEVELOPMENT AND/OR MAINTENANCE OF THE
CC NERVOUS SYSTEM. THIS IS A TYROSINE-PROTEIN KINASE RECEPTOR. KNOWN
CC SUBSTRATES FOR THE TRK RECEPTORS ARE SHC, PI-3 KINASE, AND PLC-
CC GAMMA-1.
CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC PROTEIN TYROSINE PHOSPHATE.
CC -1- SUBUNIT: EXISTS IN A DYNAMIC EQUILIBRIUM BETWEEN MONOMERIC (LOW
CC AFFINITY) AND DIMERIC (HIGH AFFINITY) STRUCTURES (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- PTM: LIGAND-MEDIATED AUTO-PHOSPHORYLATION.
CC -1- ALTERNATIVE PRODUCTS: A TRUNCATED NON-CATALYTIC FORM WHICH
CC FUNCTION REMAINS STILL UNCLEAR IS PRODUCED BY ALTERNATIVE
CC SPLICING. THE RATIO OF FULL-LENGTH TO TRUNCATED FORM IS HIGHER IN
CC FETAL THAN IN ADULT BRAIN.
CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED, MAINLY IN THE NERVOUS
CC TISSUE. IN THE CNS, EXPRESSION IS OBSERVED IN THE CEREBRAL CORTEX,
CC HIPPOCAMPUS, THALAMUS, CHOROID PLEXUS, GRANULAR LAYER OF THE
CC CEREBELLUM, BRAIN STEM, AND SPINAL CORD. IN THE PERIPHERAL NERVOUS
CC SYSTEM, IT IS EXPRESSED IN MANY CRANIAL GANGLIA, THE OPTALMIC
CC NERVE, THE VESTIBULAR SYSTEM, MULTIPLE FACIAL STRUCTURES, THE
CC SUBMAXILLARY GLANDS, AND DORSAL ROOT GANGLIA.
CC -1- SIMILARITY: BELONGS TO THE INSULIN RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES. HIGHEST SIMILARITY WITH OTHER TRK-TYPE RECEPTORS.
CC -1- SIMILARITY: THE REPEATED LEUCINE-RICH (LRR) SEGMENT IS FOUND IN
CC MANY PROTEINS. NUMBER IN THIS PROTEIN: 2.
CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CONTAINS
CC TWO C2-LIKE DOMAIN

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial/
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
CC or send an email to license@isb-sib.ch.

DR EMBL; U12140; G530791; -;
DR EMBL; S76473; G913718; -;
DR EMBL; S76474; G913720; -;
DR EMBL; X75958; G473008; -;
DR MIM; 600456; -;
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.
DR PROSITE; PS50011; PROSITE_KINASE_DOM; 1.
DR PFAM; PF00047; Ig; 1.
DR PFAM; PF00069; pkinase; 1.
DR PFAM; PF00560; LRR; 1.
DR HSP; P11362; Irgi.
DR TRANSFERASE; TYROSINE-PROTEIN KINASE; TRANSMEMBRANE; ATP-BINDING;
DR PHOSPHORYLATION; RECEPTOR; GLYCOPROTEIN; NEUROGENESIS; SIGNAL;
DR LEUCINE-REPEAT; REPEAT; IMMUNOGLOBULIN FOLD; ALTERNATIVE SPLICING.
FT SIGNAL 1 31

15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
HIGH AFFINITY NERVE GROWTH FACTOR RECEPTOR PRECURSOR (EC 2.7.1.112)
(TRK1 TRANSFORMING TYROSINE KINASE PROTEIN) (P140-TRKA) (TRK-A).
NTRK1 OR TRK.
GN HOMO SAPIENS (HUMAN).
OS EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
OC
RN SEQUENCE FROM N.A. (TRKA-I ISOFORM).
RP TISSUE=COLON;
RC MEDLINE; 89181575.
RA MARTIN-ZANCA D., OSKAM R., MITRA G., COPELAND T.D., BARBACID M.;
RT "Molecular and biochemical characterization of the human trk proto-
oncogene.";
RL MOL. CELL. BIOL. 9:24-33(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE; 95123473.
RA SHELTON D.L., SUTHERLAND J., GRIPP J., CAMERATO T., ARMANINI M.P.,
RA PHILLIPS H.S., CARROLL K., SPENCER S.D., LEVINSON A.D.;
RT "Human trks: molecular cloning, tissue distribution, and expression
of extracellular domain immunoadhesins.";
RL J. NEUROSCI. 15:477-491(1995).
RN [3]
RP SEQUENCE OF 399-796 FROM N.A.
RX MEDLINE; 86146854.
RA MARTIN-ZANCA D., HUGHES S.H., BARBACID M.;
RT "A human oncogene formed by the fusion of truncated tropomyosin and
protein tyrosine kinase sequences.";
RL NATURE 319:743-748(1986).
RN [4]
RP SEQUENCE OF 399-796 FROM N.A.
RX MEDLINE; 88196074.
RA KOZMA S.C., REDMOND S.M.S., SAURER S.M., GRONER B., HYNES N.E.;
RT "Activation of the receptor kinase domain of the trk oncogene by
recombination with two different cellular sequences.";
RL EMBO J. 7:147-154(1988).
RN [5]
RP FUNCTION.
RX MEDLINE; 91218846.
RA HEMPSTEAD B.L., MARTIN-ZANCA D., KAPLAN D.R., PARADA L.F., CHAO M.V.;
RT "High-affinity NGF binding requires coexpression of the trk proto-
oncogene and the low-affinity NGF receptor.";
RL NATURE 350:678-683(1991).
RN [6]
RP FUNCTION.
RX MEDLINE; 91191557.
RA KLEIN R., JING S., NANDURI V., O'ROURKE E., BARBACID M.;
RT "The trk proto-oncogene encodes a receptor for nerve growth factor.";
RL CELL 65:189-197(1991).
RN [7]
RP ALTERNATIVE SPLICING.
RX MEDLINE; 93315496.
RA BARKER P.A., LOMEN-HOERTH C., GENSCHE E.M., MEAKIN S.O., GLASS D.J.,
RA SHOOTER E.M.;
RT "Tissue-specific alternative splicing generates two isoforms of the
trk receptor.";
RL J. BIOL. CHEM. 268:15150-15157(1993).
RN [8]
RP VARIANT CIPA ARG-577.
RX MEDLINE; 96331294.
RA INDO Y., TSURUTA M., HAYASHIDA Y., KARIM M.A., OHTA K., KAWANO T.,
RA MITSUBUCHI H., TONOKI H., AWAYA Y., MATSUDA I.;
RT "Mutations in the TRKA/NGF receptor gene in patients with congenital
RT insensitivity to pain with anhidrosis.";
RL NAT. GENET. 13:485-488(1996).
RN [9]
RP MUTAGENESIS OF TYR-791.
RX MEDLINE; 94179299.
RA LOEB D.M., STEPHENS R.M., COPELAND T.D., KAPLAN D.R., GREENE L.A.;
RT "A trk nerve growth factor (NGF) receptor point mutation affecting
interaction with phospholipase C-gamma 1 abolishes NGF-promoted
peripherin induction but not neurite outgrowth.";
J. BIOL. CHEM. 269:8901-8910(1994).
RN [10]
RP MUTAGENESIS, AND PHOSPHORYLATION SITES.
RX MEDLINE; 94206546.
RA STEPHENS R.M., LOEB D.M., COPELAND T.D., PAMSON T., GREENE L.A.,
RA KAPLAN D.R.;
RT "Trk receptors use redundant signal transduction pathways involving
SHC and PLC-gamma 1 to mediate NGF responses.";
RL NEURON 12:691-705(1994).
RN [11]
RP STRUCTURE BY NMR OF 489-500.
RX MEDLINE; 96097066.
RA ZHOU M.M., RAVICHANDRAN K.S., OLEJNICZAK E.F., PETROS A.M.,
RA MEADOWS R.P., SATTLER M., HARLAN J.E., WADE W.S., BURAKOFF S.J.,
RA PESIK S.W.;
RT "Structure and ligand recognition of the phosphotyrosine binding
domain of Shc.";
RL NATURE 378:584-592(1995).
RN [12]
RP FUNCTION: REQUIRED FOR HIGH-AFFINITY BINDING TO NERVE GROWTH
FACTOR (NGF). NEUROTROPHIN-3 AND NEUROTROPHIN-4/5 BUT NOT BRAIN-
DERIVED NEUROTROPHIC FACTOR (BDNF). KNOWN SUBSTRATES FOR THE TRK
RECEPTORS ARE SHC, PI-3 KINASE, AND PLC-GAMMA-1. HAS A CRUCIAL
ROLE IN THE DEVELOPMENT AND FUNCTION OF THE NOCICEPTIVE RECEPTION
SYSTEM AS WELL AS ESTABLISHMENT OF THERMAL REGULATION VIA
SWEATING. ACTIVATES ERK1 BY EITHER SHC- OR PLC-GAMMA-1-DEPENDENT
SIGNALING PATHWAY.
CC CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
PROTEIN TYROSINE PHOSPHATE.
CC SUBUNIT: EXISTS IN A DYNAMIC EQUILIBRIUM BETWEEN MONOMERIC (LOW
AFFINITY) AND DIMERIC (HIGH AFFINITY) STRUCTURES.
CC SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC PTM: LIGAND-MEDIATED AUTO-PHOSPHORYLATION.
CC ALTERNATIVE PRODUCTS: TWO ISOFORMS, TRKA-I AND TRKA-II (SHOWN
HERE), ARE PRODUCED BY ALTERNATIVE SPLICING. BOTH ISOFORMS HAVE
SIMILAR BIOLOGICAL PROPERTIES.
CC TISSUE SPECIFICITY: TRKA-II IS PRIMARILY EXPRESSED IN NEURONAL
CELLS; TRKA-I IS FOUND IN NON-NEURONAL TISSUES.
CC DISEASE: THE AMINO END OF SEVERAL DIFFERENT CELLULAR PROTEINS
(SUCH AS NON-MUSCLE TROPOMYOSIN TM30NM) HAS BEEN FOUND FUSED WITH
THE PROTEIN TYROSINE KINASE DOMAIN ENCODED BY NTRK1. THESE SOMATIC
REARRANGEMENT CREATES AN ONCOGENIC PROTEIN.
CC DISEASE: DEFECTS IN NTRK1 ARE A CAUSE OF A DISEASE CALLED CIPA.
CC WHICH IS CHARACTERIZED BY A CONGENITAL INSENSITIVITY TO PAIN AND
BY ANHIDROSIS (ABSENCE OF SWEATING).
CC SIMILARITY: BELONGS TO THE INSULIN RECEPTOR FAMILY OF TYROSINE-
PROTEIN KINASES. HIGHEST SIMILARITY WITH OTHER TRK-TYPE RECEPTORS.
CC SIMILARITY: THE REPEATED LEUCINE-RICH (LRR) SEGMENT IS FOUND IN
MANY PROTEINS. NUMBER IN THIS PROTEIN: 2.
CC SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CONTAINS
TWO C2-LIKE DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; M23102; G339918; --
CC EMBL; X03541; G37403; ALT_INIT.
CC EMBL; X06704; G37400; ALT_INIT.
CC PIR; A30124; TVHUTT.
CC PDB; 1SHC; 15-MAY-97.
CC MIM; 191315; --
CC MIM; 164970; --
CC MIM; 256800; --
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
CC PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
CC PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.

DR PFAM: PF00069; pkinase; 1.
 KW TRANSFERASE; TYROSINE-PROTEIN KINASE; ATP-BINDING; PHOSPHORYLATION;
 KW RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL; REPEAT; LEUCINE-REPEAT;
 KW PROTO-ONCOGENE; CHROMOSOMAL TRANSLOCATION; NEUROGENESIS; 3D-STRUCTURE;
 KW IMMUNOGLOBULIN FOLD; DISEASE MUTATION; ALTERNATIVE SPLICING.
 FT SIGNAL 1 32
 FT CHAIN 33 796
 FT
 FT DOMAIN 33 423
 FT TRANSMEM 424 439
 FT DOMAIN 440 796
 FT CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 72 117
 FT REPEAT 72 93
 FT REPEAT 96 117
 FT REPEAT 117 117
 FT DOMAIN 211 269
 FT DOMAIN 299 365
 FT DOMAIN 510 781
 FT NP_BIND 516 524
 FT BINDING 544 544
 FT ACT_SITE 650 650
 FT MOD_RES 496 496
 FT MOD_RES 676 676
 FT MOD_RES 680 680
 FT MOD_RES 681 681
 FT MOD_RES 791 791
 FT SITE 392 393
 FT
 FT SITE 496
 FT SITE 791 791
 FT MUTAGEN 496 496
 FT MUTAGEN 544 544
 FT MUTAGEN 791 791
 FT
 FT CARBOHYD 67 67
 FT CARBOHYD 95 95
 FT CARBOHYD 121 121
 FT CARBOHYD 188 188
 FT CARBOHYD 202 202
 FT CARBOHYD 253 253
 FT CARBOHYD 262 262
 FT CARBOHYD 281 281
 FT CARBOHYD 318 318
 FT CARBOHYD 323 323
 FT CARBOHYD 338 338
 FT CARBOHYD 358 358
 FT CARBOHYD 401 401
 FT VARSPLIC 393 398
 FT VARIANT 577 577
 FT CONFLICT 300 300
 Query Match 13.3%; Score 654.5; DB 1; Length 796;
 Best Local Similarity 30.3%; Pred. No. 9.5e-27;
 Matches 185; Conservative 73; Mismatches 151; Indels 201; Gaps 19;
 QY 349 RFLPAGPMLFSEISFI-----SDVNSSPALGGTGP----- 382
 DB 314 RFLNFGS--VLNFSFIETFELEPAANETVVRGCLRLNQPTHVNGNYTLIAANPFGQAS 371
 QY 382 -----PAPWMPGPPPTNFSSLEPRGQOPVAKPEGSPTAILGCLVAIILL 431
 DB 372 ASIMAAFMDFEENPEDIPVSPVDNTSGDPEVKEDTDPFGVAVGLAVFACLF 431
 QY 432 LIALMLRLHWRLLSKAER-----VLSEE-----LTVHLSVPGDTILNNRPGP 478
 DB 432 LSTLLL-----VLNCKGRNFGINRPAVLAPEDGLAMSLHMTLGSSL----- 477
 QY 479 REPPYQEPGRGNPHPCVPNGSAL--LLSNPAYRLLLATYARPGPGPTPAWA 535
 DB 477 -----SPTEGSGGLQHIIENPOY----- 497
 QY 536 KPTNTQAYSGDMEPEKFGAPLLPPPPONSVPHYAEADIVTLQGVGTGNTYAVPALPPGA 595

DB 497FS-----DACHVHHIKRDI----- 512
 QY 596 VGDGPPRVDFPRSLRFRKEKLGEGQFGEVHLCEDVS---PQDLVSLDFPLNVRKGHPLLV 652
 DB 512 -----LKWELGEGAFCKVFLAECHNLLPEQ-----KMLV 541
 QY 653 AVKILRPDANKARNDFLKEVKIMSRUKDNIRLLGVCVQDDPLCMITDMYMGDLNQF 712
 DB 542 AVKALK-EASESARQDFORAEALLTMLQHQHIVRFEGVCTEGRPLLMVFYMRHGLNRF 600
 QY 713 LSAHQLEKAAEGAPGQQAQGTISYPMLLHVAQAIOASSMYRLATLNEVHRLATRNC 772
 DB 601 LRSRGPDAKLLAGE---DVAPGP-IGLGOLLAVASOVAAGMVLGLHFEVHDLATRNC 656
 QY 773 LVGENFTIKIADFGMSNLVAGDYRVQGRAVLPIRMMAWECILMGKFTTASDVAFGVT 832
 DB 657 LVGQGLVVKIGDFGMSRDIYSTDYRVVGRGTMLPIRMWPPESILYRKFTTESDVWSEGVV 716
 QY 833 LWEVLMCRAQPGCOLTDEQVIENAGEFFRDQGRQVYLSRPPACOGLYELMLRCWSRES 892
 DB 717 LWEFTYTK-OPWYQLSNTAIDCI-----TOGRE--LERPRACPEVYAIMRGWQREP 768
 QY 893 EQRPPFSQLH 902
 DB 769 QQRHSIKDVH 778
 RESULT 11
 TRKA_RAT
 ID TRKA_RAT STANDARD; PRT; 799 AA.
 AC P35739;
 DT 01-JUN-1994 (REL. 29, CREATED)
 DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
 DE HIGH AFFINITY NERVE GROWTH FACTOR RECEPTOR PRECURSOR (EC 2.7.1.112)
 DE (P140-TRKA) (SLOW NERVE GROWTH FACTOR RECEPTOR) (TRK-A).
 GN NTRK1 OR TRKA OR TRK.
 OS RATTUS NORVEGICUS (RAT).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.
 [1]
 RN SEQUENCE FROM N.A. (TRKA-II ISOFORM).
 RX MEDLINE; 92196121.
 RA MEKIN S.O., SUTER U., DRINKWATER C.C., WELCHER A.A., SHOOTER E.M.;
 RT "The rat trk protooncogene product exhibits properties characteristic
 RT of the slow nerve growth factor receptor.";
 RL PROC. NATL. ACAD. SCI. U.S.A. 89:2374-2378(1992).
 RN [2]
 RP ALTERNATIVE SPLICING.
 RX MEDLINE; 93315496.
 RA BARKER P.A., LOMEN-HOERTH C., GENSCHE E.M., MEKIN S.O., GLASS D.J.,
 RA SHOOTER E.M.;
 RT "Tissue-specific alternative splicing generates two isoforms of the
 RT trkA receptor.";
 RL J. BIOL. CHEM. 268:15150-15157(1993).
 CC -!- FUNCTION: REQUIRED FOR HIGH-AFFINITY BINDING TO NERVE GROWTH
 CC FACTOR (NGF). NEUROTROPHIN-3 AND NEUROTROPHIN-4/5 BUT NOT BRAIN-
 CC DERIVED NEUROTROPHIC FACTOR (BDNF). KNOWN SUBSTRATES FOR THE TRK
 CC RECEPTORS ARE SHC, PI-3 KINASE, AND PLC-GAMMA-1. HAS A CRUCIAL
 CC ROLE IN THE DEVELOPMENT AND FUNCTION OF THE NOCICEPTIVE RECEPTION
 CC SYSTEM AS WELL AS ESTABLISHMENT OF THERMAL REGULATION VIA
 CC SWATING. ACTIVATES ERK1 BY EITHER SHC- OR PLC-GAMMA-1-DEPENDENT
 CC SIGNALING PATHWAY (BY SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
 CC PROTEIN TYROSINE PHOSPHATE.
 CC -!- SUBUNIT: EXISTS IN A DYNAMIC EQUILIBRIUM BETWEEN MONOMERIC (LOW
 CC AFFINITY) AND DIMERIC (HIGH AFFINITY) STRUCTURES.
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -!- PTM: LIGAND-MEDIATED AUTO-PHOSPHORYLATION.
 CC -!- ALTERNATIVE PRODUCTS: TWO ISOFORMS, TRKA-I AND TRKA-II (SHOWN
 CC HERE). ARE PRODUCED BY ALTERNATIVE SPLICING. BOTH ISOFORMS HAVE
 CC SIMILAR BIOLOGICAL PROPERTIES.

CC -1- TISSUE SPECIFICITY: TRKA-II IS PRIMARILY EXPRESSED IN NEURONAL
CC CELLS; TRKA-I IS FOUND IN NON-NEURONAL TISSUES.
CC -1- SIMILARITY: BELONGS TO THE INSULIN RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES. HIGHEST SIMILARITY WITH OTHER TRK-TYPE RECEPTORS.
CC -1- SIMILARITY: THE REPEATED LEUCINE-RICH (LRR) SEGMENT IS FOUND IN
CC MANY PROTEINS. NUMBER IN THIS PROTEIN: 2.
CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CONTAINS
CC TWO C2-LIKE DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: M85214; G207482;
CC EMBL: L12225; NOT_ANNOTATED_CDS.
CC PIR: A41981; TVRTTB.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00239; RECEPTOR_TYR_KIN_II; 1.
DR PFM: PF00069; pkinase; 1.
DR PFM: PF00560; LRR; 2.
DR HSP: P11362; LRG1.
KW TRANSFERASE; TYROSINE-PROTEIN KINASE; ATP-BINDING; PHOSPHORYLATION;
KW RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL; REPEAT; LEUCINE-REPEAT;
KW IMMUNOGLOBULIN FOLD; NEUROGENESIS; ALTERNATIVE SPLICING.
FT SIGNAL 1 32
FT CHAIN 33 799
FT DOMAIN 33 418
FT TRANSMEM 419 442
FT DOMAIN 443 799
FT DOMAIN 72 117
FT REPEAT 72 93
FT REPEAT 96 117
FT DOMAIN 213 271
FT DOMAIN 301 368
FT DOMAIN 513 784
FT NP_BIND 519 527
FT BINDING 547 547
FT ACT_SITE 553 653
FT MOD_RES 439 499
FT MOD_RES 679 679
FT MOD_RES 683 683
FT MOD_RES 684 684
FT MOD_RES 794 794
FT SITE 499 499
FT SITE 794 794

FT CARBOHYD 67
FT CARBOHYD 121 121
FT CARBOHYD 190 190
FT CARBOHYD 204 204
FT CARBOHYD 255 255
FT CARBOHYD 264 284
FT CARBOHYD 320 320
FT CARBOHYD 325 325
FT CARBOHYD 341 341
FT CARBOHYD 361 361
FT CARBOHYD 404 404
FT VARSPIC 396 401
SQ SEQUENCE 799 AA; 87868 MW; FE1B402 CRC32;
Query Match 13.3%; Score 653.5; DB 1; Length 799;
Best Local Similarity 29.8%; Pred. No. 1.1e-26;
Matches 190; Conservative 78; Mismatches 155; Indels 215; Gaps 21;

QY 335 SVPLGRVARFLOC-----RFLPAGPMLLSEISFI----- 366
Db 289 SVHLKAVEQHHWCIPFSVDGPAPSLRWFFNGS--VLNETSFTTQFLESALTNETMRH 346
QY 366 -----SDVNNSSPALGGTFP-----PAPWPPGPPPTNFSSILEPR 403
Db 347 GCLRNLQPTHVNGNYTLAANPYGQAAASIAAFPMNPFFNPEDPIPVSFSPVDINST 406
QY 404 GOOPVAKPEGSPTAILGCLVAITLIIILLLLMLWRLWRLRLSKAERR-----V 454
Db 407 SDPVEKDETFPGVSVAVLAVSAALFL-SALLL-----VLNKGQSRKFGINRAV 458
QY 455 LLEE-----LTVHLSVPGDTILINNRGPPPPYOEPRGNPPHSAFCVNGSAL---L 507
Db 459 LAPEDGLAMSLHFTLGGSSL-----SPTGKGSGLOGHI 493
QY 508 LSNPAYRLLLATYARPPRGPGPTPAWAKPTNQAYSQDYMPEKPGAPLLPPPPQNSVP 567
Db 494 MENPQY-----FS-----DTCVH 506
QY 568 HYAEADIVTLQGTGNTYAVPALPPGAVGDPGRVDFPRSLRKEKLGEGQGEVHLC 627
Db 507 HIKRQDII-----LKWELGEGAFKGVFLA 530
QY 628 E---VDSPODLVSLDFPLNVRKGHPLLVAVKLLRPDANKNARNDLFKEVKIMSRKLDNI 684
Db 531 ECVNLLNDQD-----KMLVAVKALK-ETSENARQDFHREAELTLMLOHQHI 575
QY 685 IRLGVCVODDPLCMITDYMGNDLQFLSAHQLEDKAAEGAPDQGAAGPTISYPMLL 744
Db 576 VRFVGCTGGGLLVMFEYMRHGDNLNRLRSHGPDAKLLAGGE---DVAPGP-LGLGQLL 631
QY 745 HVAQAISGMRYLATLNFVHRDLATRNCLVGNFTIKIADFGMSRLNAGDYRYVOGRAV 804
Db 632 AVASQVAAGVYLASLHFVHRDLATRNCLVGLVYKIGDFGMSRDIYDYRVGGRTM 691
QY 805 LPIRWANECILMGKTTASDVWAGVTLWEVLMCLRAQPFQGLTDEQVIENAGEFFRDO 864
Db 692 LPIRWMPPEISILYRKFTESDVMSFGVWLVEITFYGK-OPWYQLSNTEAIECI----TQ 745
QY 865 GROVILSRPACPGQGLYELMRCWSRESQRPFPFSOLH 902
Db 746 GRE--LERPRACPDVYAIMRGQWQREPOQLRSMKRDVH 781
RESULT 12
TRKA_CHICK
ID TRKA_CHICK STANDARD; PRT; 778 AA.
AC Q91009; 1998 (REL. 36, CREATED)
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE HIGH AFFINITY NERVE GROWTH FACTOR RECEPTOR PRECURSOR (EC 2.7.1.112)
DE (TRK-A) (FRAGMENT).
GN TRKA.
OS GALLUS GALLUS (CHICKEN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ARCHOSAURIA; AVES;
OC NEOGNATHAE; GALLIFORMES; PHASIANIDAE; PHASIANINAE; GALLUS.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-SYMPATHETIC GANGLION;
RX MEDLINE; 97047187.
RA BACKSTROM A., SODERSTROM S., KYLBERG A., EBENDAL T.;
RT "Molecular cloning of the chicken trka and its expression in early
J. NEUROSCI. RES. 46:67-81(1996)."
CC -1- FUNCTION: REQUIRED FOR HIGH-AFFINITY BINDING TO NERVE GROWTH
FACTOR (NGF), NEUTROPHIN-3 AND NEUTROPHIN-4/5 BUT NOT BRAIN-
DERIVED NEUROTROPHIC FACTOR (BDNF). KNOWN SUBSTRATES FOR THE TRK
RECEPTORS ARE SHC, PI-3 KINASE, AND PLC-GAMMA-1 (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE - ADP +
PROTEIN TYROSINE PHOSPHATE.
CC -1- PFM: LIGAND-MEDIATED AUTO-PHOSPHORYLATION (BY SIMILARITY).

CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- SUBUNIT: EXISTS IN A DYNAMIC EQUILIBRIUM BETWEEN MONOMERIC (LOW
CC AFFINITY) AND DIMERIC (HIGH AFFINITY) STRUCTURES (BY SIMILARITY).
CC -!- DEVELOPMENTAL STAGE: EXPRESSED IN THE CONDENSING DORSAL ROOT
CC GANGLIA AT EMBRYONAL DAY 3, AND IN THE PRIMARY SYMPATHETIC CHAIN
CC GANGLIA AT EMBRYONAL DAY 4.
CC -!- ALTERNATIVE PRODUCTS: TWO ISOFORMS ARE PRODUCED BY ALTERNATIVE
CC SPLICING OF THE SAME GENE.
CC -!- SIMILARITY: BELONGS TO THE INSULIN RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES. HIGHEST SIMILARITY WITH OTHER TRK-TYPE RECEPTORS.
CC -!- SIMILARITY: THE REPEATED LEUCINE-RICH (LRR) SEGMENT IS FOUND IN
CC MANY PROTEINS. NUMBER IN THIS PROTEIN: 2.
CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CONTAINS
CC TWO C2-LIKE DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X93581; E213280; -
CC PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
CC PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
CC PROSITE: PS00239; RECEPTOR_TYR_KIN_II; 1.
CC PFAM: PF00069; pkinase; 1.
CC PFAM: PF00360; LRR; 1.
CC HSP: P11362; IGLI.
CC TRANSFERASE; TYROSINE-PROTEIN KINASE; ATP-BINDING; PHOSPHORYLATION;
CC RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN; REPEAT; LEUCINE-REPEAT;
CC IMMUNOGLOBULIN FOLD; NEUROGENESIS; SIGNAL.
CC
CC FT NON_TER 1 1
CC FT SIGNAL <1 14
CC FT CHAIN 15 778
CC
CC FT DOMAIN 15 400
CC FT TRANSMEM 401 421
CC FT DOMAIN 422 778
CC FT DOMAIN 51 96
CC FT REPEAT 51 72
CC FT REPEAT 75 96
CC FT REPEAT 193 250
CC FT DOMAIN 281 347
CC FT DOMAIN 493 763
CC FT NP_BIND 499 507
CC FT BINDING 527 527
CC FT ACT_SITE 633 633
CC FT MOD_RES 479 479
CC FT MOD_RES 659 659
CC FT MOD_RES 663 663
CC FT MOD_RES 664 664
CC FT MOD_RES 773 773
CC FT SITE 479 479
CC
CC FT SITE 773 773
CC
CC FT CARBOHYD 100 100
CC FT CARBOHYD 130 130
CC FT CARBOHYD 143 143
CC FT CARBOHYD 151 151
CC FT CARBOHYD 194 194
CC FT CARBOHYD 234 234
CC FT CARBOHYD 262 262
CC FT CARBOHYD 300 300
CC FT CARBOHYD 320 320
CC FT CARBOHYD 340 340
CC FT CARBOHYD 384 384
CC SEQUENCE 778 AA; 87338 MW; 9D26DBD0 CRC32;

Query Match 13.1%; Score 644.5; DB.1; Length 778;
Best Local Similarity 31.4%; Pred. No. 3e-26;
Matches 195; Conservative 75; Mismatches 177; Indels 175; Gaps 23;

QY 356 WLLFSEISFSDVNNSSPALGCTFPAPWPPGPPPTNFSSLELEPRGQOPVAKPEGSP 415
DB 226 WELVLEINNISSLNHK-----DLTCRAENSVGLAEDS- 259
QY 416 TAILIGLVAIIILLIIIALMLRLHWR-----RLLSKARRVLEELTVHLSV- 466
DB 259 -VMLNVTFPVILLISRAIPQHEWCIPFVSVDNPTFRILWLFNGSMLPEGPIHTRIVEY 317
QY 466 -PGDTILIN-----NRGPREPPPYQEPGRGNPHGAPCPNGS-ALLSNPARYLLLAT 519
DB 318 EPNSTVLHGCLQNR-----PTH-----VNNGNITLVQNP- LGR 351
QY 520 YAREPRG-----PGPPTPAWAKPTQAYSGDMEPEKPGAPLPPPPQNSVPHY 569
DB 352 AARSIOGRFMDNPFSPPEPIPVISPLGRNSS-----LEGPVETADEHT 398
QY 570 AEADIVTLQGV-----TGGNTYAVPALPPG-----AVGDGP 600
DB 399 FGVSVAVALAFASFLSVMLIALNKGHSKFGINSRAVLAPEDGLAMSLHFTLGGSP 458
QY 601 -----PRVD-----PPRSRLRFEKLGEGQGEVHLCVDS--PQ 633
DB 459 VSSPESKLDGLKSNFIENPOYFCNACVHHVQRRDIIVLKWELGEGAFKVFLEACSHLLPE 518
QY 634 DLVSLDFPLNVRGHPLLVAVKILRPDATNARNDELKEVKIMSRKLDKPNILLLGVCVQ 693
DB 519 Q-----EKTLLVAVKALK-EVTENARLDFOREAELLTVLQHEHIVAFYGVCTE 564
QY 694 DDPLCMITDMYNGDLNQFLSAHOLEDKAAEGAPGCGQAAAGPTISYPMLLHVAAQIASG 753
DB 565 GDPLIMVFEYMKHGDNLNRLSHGPDAKILD--QGQPCGQGLTSLH--MLQIATQIASG 620
QY 754 MRYLATLNFVHRDLATRNCLVNGENTIKIADFGMSRNLVAGDYRYVROGRAVLPIRMWAE 813
DB 621 MVLASLHFVHRDLATRNCLVGHDLVVKIGDFGMSRDIYSTDIYRVGRTMLPIRMWPE 680
QY 814 CILMGKFTTASDWAFVGTLLWEVLMCLCRAOPFGOLTDEQVLENAGEFFRDQGHQVYLSRP 873
DB 681 SILYRKFTTESDIWSFGVWLWEIPTYCK-QPWYQLSNTAEIET-----TQGRE--LRR 732
QY 874 PACPGLYELMLRCWSRESOR 895
DB 733 RTCPSEYVDIMQSCWQREPQOR 754

RESULT 13
TRKC_PIG STANDARD; PRT; 825 AA.
ID TRKC_PIG
AC P24786; 1992 (REL. 21, CREATED)
DT 01-MAR-1992 (REL. 21, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE NT-3 GROWTH FACTOR RECEPTOR PRECURSOR (EC 2.7.1.112) (TRKC TYROSINE
DE KINASE) (GPI45-TRKC) (TRK-C).
GN NTRK3 OR TRKC.
OS SUS SCROFA (FIG).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC ARTIODACTYLA; SUIFORMES; SUINA; SUIDAE; SUS.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN:
RX MEDLINE; 91364178.
RT LAMBALLE F., KLEIN R., BARBACID M.;
RT "trkc, a new member of the trk family of tyrosine protein kinases, is
RT a receptor for neurotrophin-3."
RL CELL 66:967-979(1991).
CC -!- FUNCTION: RECEPTOR FOR NEUROTROPHIN-3 (NT-3). THIS IS A TYROSINE-
CC PROTEIN KINASE RECEPTOR. KNOWN SUBSTRATES FOR THE TRK RECEPTORS
CC ARE SHC, PI-3 KINASE, AND PLC-GAMMA-1.

15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
 NT-3 GROWTH FACTOR RECEPTOR PRECURSOR (EC 2.7.1.112) (TRKC TYROSINE KINASE) (TRK-C).
 GN. TRK.
 OS. GALLUS GALLUS (CHICKEN).
 OC. EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ARCHOSAURIA; AVES;
 OC. NEOGNATHAE; GALLIFORMES; PHASIANIDAE; PHASIANINAE; GALLUS.
 RN. [1]
 RP. SEQUENCE FROM N.A.
 RX. MEDLINE: 94338700.
 RA. GARNER A.S., LARGE T.H.;
 RA. "Isoforms of the avian trkC receptor: a novel kinase insertion
 RT. dissociates transformation and process outgrowth from survival.";
 RL. NEURON 13:457-472(1994).
 RN. [2]
 RP. SEQUENCE OF 378-513 FROM N.A.
 RX. TISSUE-EMBRYO;
 RA. MEDLINE: 94084905.
 RA. WILLIAMS R., BACKSTROM A., EBENDAL T., HALLBOOK F.;
 RT. "Molecular cloning and cellular localization of trkC in the chicken
 RL. embryo.";
 RN. BRAIN RES. 75:235-252(1993).
 CC. -1- FUNCTION: RECEPTOR FOR NEUROTROPHIN-3 (NT-3). THIS IS A TYROSINE-
 CC. PROTEIN KINASE RECEPTOR. KNOWN SUBSTRATES FOR THE TRK RECEPTORS
 CC. ARE SHC, PI-3 KINASE, AND PLC-GAMMA-1.
 CC. -1- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
 CC. PROTEIN TYROSINE PHOSPHATE.
 CC. -1- PTM: LIGAND-MEDIATED AUTO-PHOSPHORYLATION.
 CC. -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC. -1- SUBUNIT: EXISTS IN A DYNAMIC EQUILIBRIUM BETWEEN MONOMERIC (LOW
 CC. AFFINITY) AND DIMERIC (HIGH AFFINITY) STRUCTURES.
 CC. -1- DEVELOPMENTAL STAGE: EXPRESSION OCCURS IN THE EMBRYONAL DAY 2 (E2)
 CC. EMBRYO WITH INCREASING LEVELS LATER IN DEVELOPMENT. IN THE E9
 CC. EMBRYO HIGHEST LEVELS ARE FOUND IN BRAIN AND SPINAL CORD WITH
 CC. INTERMEDIATE LEVELS IN EYE, HEART, GUT AND MUSCLE. LOW LEVELS ARE
 CC. FOUND IN KIDNEY, LIVER, SKIN AND YOLK SAC.
 CC. -1- ALTERNATIVE PRODUCTS: A TRUNCATED FORM IS PRODUCED BY ALTERNATIVE
 CC. SPLICING. THIS FORM IS PREFERENTIALLY EXPRESSED IN THE ADULT
 CC. CENTRAL NERVOUS SYSTEM, BUT SCARCELY DURING THE EMBRYONIC STAGES.
 CC. -1- SIMILARITY: BELONGS TO THE INSULIN RECEPTOR FAMILY OF TYROSINE-
 CC. PROTEIN KINASES. HIGHEST SIMILARITY WITH OTHER TRK-TYPE RECEPTORS.
 CC. -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CONTAINS
 CC. TWO C2-LIKE DOMAIN.
 CC. -----
 CC. This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC. between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC. the European Bioinformatics Institute. There are no restrictions on its
 CC. use by non-profit institutions as long as its content is in no way
 CC. modified and this statement is not removed. Usage by and for commercial
 CC. entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC. or send an email to license@isb-sib.ch).
 CC. -----
 DR. EMBL: S74248; G712821;
 DR. EMBL: Z30091; E53026;
 DR. PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR. PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 DR. PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR. PROSITE: PS00239; RECEPTOR_TYR_KIN_II; 1.
 DR. PFAM: PF00047; Ig; 1.
 DR. PFAM: PF00069; pkinase; 1;
 DR. PFAM: PF00560; LRR; 1.
 DR. HSSP: P11362; IFGI.
 DR. TRANSFERASE; TYROSINE-PROTEIN KINASE; TRANSMEMBRANE; ATP-BINDING;
 KW. PHOSPHORYLATION; RECEPTOR; GLYCOPROTEIN; NEUROGENESIS; SIGNAL;
 KW. ALTERNATIVE SPLICING; IMMUNOGLOBULIN FOLD.
 FT. SIGNAL
 FT. CHAIN 1 * 31 BY SIMILARITY.
 FT. DOMAIN 32 852 NT-3 GROWTH FACTOR RECEPTOR.
 FT. DOMAIN 32 430 EXTRACELLULAR (POTENTIAL).
 FT. TRANSMEM 431 455 POTENTIAL.
 FT. DOMAIN 456 852 CYTOPLASMIC (POTENTIAL).
 FT. DOMAIN 227 288 IG-LIKE C2-TYPE DOMAIN.
 FT. DOMAIN 219 382 IG-LIKE C2-TYPE DOMAIN.
 FT. DOMAIN 540 841 PROTEIN KINASE.

FT NP_BIND	546	554	ATP (BY SIMILARITY).
FT BINDING	574	574	ATP (BY SIMILARITY).
FT ACT_SITE	681	681	BY SIMILARITY.
FT MOD_RES	518	518	PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES	707	707	PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES	711	711	PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES	712	712	PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES	847	847	PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT SITE	518	518	INTERACTION WITH SHC PROTEIN (BY SIMILARITY).
FT SITE	847	847	INTERACTION WITH PLC-GAMMA1 (BY SIMILARITY).
FT CARBOHYD	68	68	POTENTIAL.
FT CARBOHYD	72	72	POTENTIAL.
FT CARBOHYD	79	79	POTENTIAL.
FT CARBOHYD	163	163	POTENTIAL.
FT CARBOHYD	203	203	POTENTIAL.
FT CARBOHYD	218	218	POTENTIAL.
FT CARBOHYD	232	232	POTENTIAL.
FT CARBOHYD	259	259	POTENTIAL.
FT CARBOHYD	267	267	POTENTIAL.
FT CARBOHYD	272	272	POTENTIAL.
FT CARBOHYD	294	294	POTENTIAL.
FT CARBOHYD	375	375	POTENTIAL.
FT CARBOHYD	388	388	POTENTIAL.
FT CONFLICT	378	378	I -> F (IN REF. 2).
SQ SEQUENCE	852 AA;	95859 MW;	3869FB1B CRC32;

Query Match 12.5%; Score 614; DB 1; Length 852;

Best Local Similarity 42.1%; Pred. No. 1.2e-24;

Matches 139; Conservative 49; Mismatches 82; Indels 60; Gaps 10;

QY	607	RSRLKFKELGSGGEGVHLCEVDS	PDQLVSLDPLNVKRGHPLLVAVKILRPD	ATKNAR	666
DB	537	RDIVLKLRELGEAGKVFLEACYNLS	-----PTNDK	-----MLVAVKALK-DPTLAAR	584

QY	667	NDFLKEVKIMSLKDPNIRILGVCVQDD	PLCMITDYMENGLNQFLSAHQLEDKAAEGA	726	
DB	585	KDFQREALLTLNQHIEHIVKFGVCGD	GPLIMVFEYMKHGDNLKFLRAH	-----G	635

QY	727	PG-----DQAAAGP-TISYPMLLHVA	QAISGRYRLATLNFVHRDLATRNCLYGENFTI	780
DB	636	PDAMILVDGPPQAKGELGSLQMLHIA	SQIASGMVYLASQHFVHRDLATRNCLVGANLLV	695

QY	781	KIADFGMSRLYAGDYR-----	VOGRAVLPLRWMAECI	815
DB	696	KIGDFGMSRDVSTDIYREGPRPKQL	STAWQRHLAPPAATVGGHTMLPIRWMPESI	755

QY	816	LMGKFTTASDVWAFVGLTWELVLMCRA	QPFQOLTDEQVFNAGFEFFRDQGRVYLSRPPA	875
DB	756	MYRKFTTESDVWVSGVILWEIFTYCK	-QPWFQLSNTEVIECI-----TQGR--VLERPRV	807

QY	876	CPQGLYELMLRCWSRESEORPFPSQL	HRFL	905
DB	808	CPKEYDMLGCMQREPOORLNKEIKYL	837	

Search completed: November 4, 1999, 01:54:22

Job time: 7597 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 4, 1999, 01:53:54 ; Search time 25.72 Seconds
(without alignments)
954.740 Million cell updates/sec

Title: US-08-170-558-8
Perfect score: 2175
Sequence: 1 DADMKGHDPKCRYALGMQ.....LELEPRGQPVAKPEGSPTA 399

Scoring table: BLOSUM62

Searched: 201082 seqs, 61543640 residues

Database : SPTREMBL_10.*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.rodent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1988.5	91.4	911	11	O35407 mus musculus
2	391	18.0	737	5	Q17305 caenorhabdi
3	381.5	17.5	737	5	Q18163 caenorhabdi
4	349	16.0	791	5	Q19354 caenorhabdi
5	222	10.2	931	4	O60462 homo sapien
6	222	10.2	909	4	O14820 homo sapien
7	222	10.2	926	4	O14821 homo sapien
8	220	10.1	909	11	O35373 mus musculus
9	220	10.1	926	11	O35374 mus musculus
10	220	10.1	931	11	O35375 mus musculus
11	220	10.1	901	11	O35376 mus musculus
12	220	10.1	906	11	O35377 mus musculus
13	220	10.1	914	11	O35378 mus musculus
14	217	10.0	925	11	O35276 mus musculus
15	211.5	9.7	845	4	Q14113 homo sapien
16	204	9.4	923	4	O60461 homo sapien
17	204	9.4	923	4	O14786 homo sapien
18	201.5	9.3	1128	11	O88442 mus musculus
19	198.5	9.1	784	11	O54860 mus musculus
20	198.5	9.1	722	11	O92100 mus musculus
21	192	8.8	216	4	Q14286 homo sapien
22	189.5	8.7	2183	11	O88783 mus musculus
23	184.5	8.5	558	5	O76470 lytechinus
24	184	8.5	2343	6	O18806 canis famil
25	182	8.4	2343	6	O62730 canis famil
26	179.5	8.3	2224	4	O43737 homo sapien
27	174	8.0	480	11	O35474 mus musculus
28	172.5	7.9	728	6	O97567 bos taurus
29	167	7.7	480	4	O43854 homo sapien

30	167	7.7	224	4	O15537	O15537 homo sapien
31	164.5	7.6	719	11	Q61281	Q61281 mus musculus
32	163	7.5	224	11	Q921L4	Q921L4 mus musculus
33	150	6.9	1384	4	P78357	P78357 homo sapien
34	148.5	6.8	1283	5	Q94887	Q94887 drosophila
35	135	6.2	1385	11	O54991	O54991 mus musculus
36	135	6.2	1381	11	P97846	P97846 rattus norv
37	134.5	6.2	363	6	O77718	O77718 equus cabal
38	106.5	4.9	1008	13	Q9YHV5	Q9YHV5 fugu rubrip
39	105.5	4.9	4123	4	O75851	O75851 homo sapien
40	97	4.5	1101	5	O27235	O27235 drosophila
41	97	4.5	892	5	O27436	O27436 drosophila
42	97	4.5	1269	5	Q05817	Q05817 drosophila
43	94.5	4.3	395	2	P74597	P74597 synechocyst
44	93	4.3	365	2	O927A9	O927A9 chlamydia p
45	92.5	4.3	533	11	Q9Z222	Q9Z222 mus musculus

ALIGNMENTS

RESULT 1
O35407
ID O35407 PRELIMINARY; PRT; 911 AA.
AC O35407
DT 01-JAN-1998 (TREMREL. 05, Created)
DT 01-JAN-1998 (TREMREL. 05, Last sequence update)
DT 01-MAY-1999 (TREMREL. 10, Last annotation update)
DE RECEPTOR-LIKE TYROSINE KINASE.
GN NEP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=129;
RX MEDLINE; 93390947.
RA ZERLIN M., JULIUS M.A., GOLDFARB M.;
RT "NEP: a novel receptor-like tyrosine kinase expressed in
RT proliferating neuroepithelia";
RL Oncogene 8:2731-2739(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129;
RA ZERLIN M., JULIUS M.A., GOLDFARB M.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF026259; AAB81866.1;
DR PFAM; PF00754; F5_F8_type.C; 1.
DR PFAM; PF00069; pkinase; 1.
DR PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.
DR PROSITE; PS01285; FA58C_1; 1.
DR PROSITE; PS01286; FA58C_2; 1.
SQ SEQUENCE 911 AA; 101091 MW; 6B91AABB CRC32;

Query Match		91.4%	Score 1988.5;	DB 11;	Length 911;
Best Local Similarity		91.8%;	Pred. No. 5.8e-169;		
Matches 367;		Conservative 10;	Mismatches 18;	Indels 5;	Gaps 2;
QY	1	DADMKGHDPKCRYALGMQRTIPDSISASSSSWSDSTAARHSRLSSDGDGAWCPAGS	60		
Db	20	DADMKGHDPKCRYALGMQDETIPDSISVSSSSWSDSTAARHSRLSSDGDGAWCPAGP	79		
QY	61	VFPKKEEYLVQDLQRLHLVALVGTQGRHAGGLGKEFSYRLYSRDGRRWMDKRWGQ	120		
Db	80	VFPKKEEYLVQDLRLHLVALVGTQGRHAGGLGKEFSYRLYSRDGRRWMDKRWGQ	139		
QY	121	EVISGNEDPEGVWLKDLGPPVRLVREYPRADRVMSVCLRVLYGCLWRDGLLSYAPV	180		
Db	140	EVISGNEDPGVWLKDLGPPVRLVREYPRADRVMSVCLRVLYGCLWRDGLLSYAPV	199		
QY	181	GOTMYLSEA-VYLDNSTYDGHVTGVLQYGLGLQADGVVGLDDEFKQQLRWPGDYVG	239		

Db 200 GQTMQLSEVMVHLNDSTYDGTAGGLQYGLGLQADGWGLDDFRSQELRVWPGYDVG 259
 Qy 240 WSNISFSGYVEMFEEDRLRAFQAMQVHCNNMHTLGLARLPGGVECFRSGPAMWEGEP 299
 Db 260 WSNOSFTGYVEMFEEDRLTFQTMQVHCNNMHTLGLARLPGGVECFRSGPAMWEGEP 319
 Qy 300 MRHLNGLGDPARAVSVPLGGVAVRFLOCRFLFAGPWLFLFSISFISDVWNSSPALG 359
 Db 320 VRHALLGSLGDPARATSVPLGGVHGRFLQCRFLFAGPWLFLFSISFISDVWNS 376
 Qy 360 GTFFPAPWMPGPPPTNFSSLELEPRGQOPVAKPEGSPTA 399
 Db 376 DTFPPAPWMPGPPPTNFSSLELEPRGQOPVKAEGSPTA 415

RESULT 2
 Q17305 PRELIMINARY; PRT; 737 AA.
 AC Q17305;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-MAY-1999 (Tremblrel. 10, Last annotation update)
 DE GO1D9.2 PROTEIN.
 GN GO1D9.2.
 OS Caenorhabditis briggsae.
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-GUJARAT G16;
 RA WU X., LE T.T.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-GUJARAT G16;
 RA WATERSTON R.;
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U56248; AAA98703.1;
 DR PFAM; PF00754; F5_F8_type_C; 1.
 DR PFAM; PF00069; pkinase.2.
 SQ SEQUENCE 737 AA; 82950 MW; 36328080 CRC32;

Query Match 18.0%; Score 391; DB 5; Length 737;
 Best Local Similarity 30.5%; Pred. No. 1e-26;
 Matches 106; Conservative 60; Mismatches 149; Indels 32; Gaps 14;
 QY 19 MODRTIPDSISASSW-SDSTAARHSRLSSDGDGAWCPAGSVFPKKEEYLOVDLQRLH 77
 Db 1 MONGDVLDQAITASSSFDKQSGVGPQNALHSELASGAWCPKQINSKSYEFLOVTLNDF 60
 QY 78 LVALVGTQGRHAGGLGKEFSRSLRYSDRRRWGKDRWGOEIVSGNEDPEGVVKDL 137
 Db 61 LITSVETQGRYNGTGREFAHYMDYMRPGSQWIRYLNRSVHYVYDGNFTTTPVIRAL 120
 QY 138 GPMWALRVFPYPRADRVMSVCLRVLYGCLWRDGLSY-TAPVGTMYLSEAVYLDST 196
 Db 121 DPPIVASRIYFPSSKNTRTVCNRAEIHGCK-HEGVYIYSTVPDGSRL---DTLDFKDSM 176
 Db 121 DPPIVASRIYFPSSKNTRTVCNRAEIHGCK-HEGVYIYSTVPDGSRLDTLDFKDNMFE 178
 QY 193 NSTYDCHTVGGLOYGGLQADGVVGLDDFRKQSELRVWPGYDYVWNSHFSFGYVE 251
 Db 179 NSQMY---TESGIT-RGLGLLTDGYVAQTSPEKQNMNSW-----IGW-NRDTTDDGRVT 228
 QY 252 MEFEEDRLRAFQAMQVHCNNMHTLGLARLPGGVECFRSGPAMWEGEPMRNLGN 307
 Db 229 ILFEFEIHNFTDVV-----LATFGNRI-DGIDVIFSDQKTFPLFSQISSSERGTLANNT 282
 QY 308 LGDPRARAVSVPLGGVAVRFLOCRFLFAGPWLFLFSISFISDVWNS 354
 Db 283 ---SRRYDFRVPLNRAGRKVRISIKFSFSSDWMFLTEVHFTSGIFKHT 326

RESULT 3
 Q19354 PRELIMINARY; PRT; 791 AA.
 ID Q19354

Q18163 PRELIMINARY; PRT; 737 AA.
 ID Q18163;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-MAY-1999 (Tremblrel. 10, Last annotation update)
 DE SIMILAR TO PROTEIN TYROSINE KINASE.
 GN C25F6.4.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 94150718.
 RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
 RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans.";
 RN Nature 368:32-38(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA BENILEY D.;
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA WATERSTON R.;
 RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U39742; AAA80432.1;
 DR PFAM; PF00754; F5_F8_type_C; 1.
 DR PFAM; PF00069; pkinase; 2.
 SQ SEQUENCE 737 AA; 83396 MW; B5F5CD5A CRC32;

Query Match 17.5%; Score 381.5; DB 5; Length 737;
 Best Local Similarity 30.3%; Pred. No. 7.3e-26;
 Matches 103; Conservative 58; Mismatches 150; Indels 29; Gaps 12;
 QY 19 MODRTIPDSISASSW-SDSTAARHSRLSSDGDGAWCPAGSVFPKKEEYLOVDLQRLH 77
 Db 1 MONGDIADSOITASSSFDKQSGVGPQNALHSELASGAWCPKQINSKSYEFLOVTLNDF 60
 QY 78 LVALVGTQGRHAGGLGKEFSRSLRYSDRRRWGKDRWGOEIVSGNEDPEGVVKDL 137
 Db 61 LITSVETQGRYNGTGREFAHYMDYMRPGSQWIRYLNRSVHYVYDGNFTTTPVIRVL 120
 QY 138 GPMWALRVFPYPRADRVMSVCLRVLYGCLWRDGLSY-TAPVGTMYLSEAVYLDST 196
 Db 121 DPPIVASRIYFPSSKNTRTVCNRAEIHGCK-HEGVYIYSTVPDGSRL---DTLDFKDSM 176
 QY 197 YDG---HTVGGL-QYGGLOLADGVVGLDDFRKQSELRVWPGYDYVWNSHFSFGYVE 252
 Db 177 FEDSQIYTESGIRKRYGWPANRRFRSAASPEKQNMNSW-----IGW-NRDTTDDGRITI 230
 QY 253 EFEDRLRAFQAMQVHCNNMHTLGLARLPGGVECFRSGPAMWEGEPMRNLGN 308
 Db 231 LFEFEIHNFTDVV-----LATFGNRI-DGIDVIFSDQKTFPLFSQISSSERQSLNNT- 284
 QY 309 GDPRARAVSVPLGGVAVRFLOCRFLFAGPWLFLFSISFIS 348
 Db 284 ---SRRYDFRVPLNRAGRKVRISIKFSFSSDWMFLTEVHFTS 321

RESULT 4
 Q19354 PRELIMINARY; PRT; 791 AA.
 ID Q19354

AC Q19354;
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DE 01-MAY-1999 (TRENBLrel. 10, Last annotation update)
 DE SIMILAR TO TYROSINE KINASE. NCBI GI: 1118035.
 GN F11D5.3
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Secernentes; Rhabditia; Rhabditidae;
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderae; Caenorhabditis.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 94150718.
 RA WILSON R., AINSWORTH R., ANDERSON K., BAYNES C., BERKS M.,
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
 RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOLDMAN P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RL Nature 368:32-38(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA NHAN M.;
 RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA WATERSTON R.;
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U41532; AAA83276.1; -;
 DR PFAM; PF00754; F5_F8_type_C; 1.
 DR PFAM; PF00629; pkinase; 1.
 SQ SEQUENCE 791 AA; 89990 MW; 991C9D67 CRC32;

Query Match 16.0%; Score 349; DB 5; Length 791;
 Best Local Similarity 25.6%; Pred. No. 6.3e-23;
 Matches 88; Conservative 67; Mismatches 143; Indels 46; Gaps 11;

QY 28 DISASSWSDTAAR-HSRLSSDGDGAWCPAGSVFPKKEEYLVQDLQRLHLVALVGTQ 86
 DB 34 DLNMLGKWFNGYSFEIRTTAHOESGSGAWCPKQINSLSKWQLQISFSDVITSVETQ 93
 QY 87 RHAGGLGKFEFSYRLVSRDQ-RRWGKWKDRGQEVISGNEDEPGVVKDLGPPVWRL 145
 DB 94 REDDGRGMEYATAFKIYWRPSLNAWASYKDDFELETIPANNNDTHAIRHLRAIARR 153
 QY 146 VFYPRADRVMSVCLRVLYGLWRDGLLSYTPVPGQMYLSEAVYLVNDSTVDGHTVGL 205
 DB 154 IRIVPVSNSTRVCHRVVEFGCPDSDLVYVNDQD---LQSGISYHDFSYDGNLANS 210
 QY 206 QY-GLGLQADGVGLDPRKQELRWPGYDVGWNSHFSFGYVMEFEFRLRAFQA 264
 DB 211 HLTGGIGKLYDGEVGNVFNH-----HWGVRKK--RNGNVKLAFFESLRLNSG 261
 QY 265 MOVHCHNMHTLCARLPGGVCEFRGRCPAMWEGPVRNINLG-----NLGDR----- 313
 DB 262 ILIHTSN-----EFKKS-AKAFSSATVLSFNGKQSDFTIVHFNPDFTES 306
 QY 313 --ARAVSVPLGGRVAFRLQCRLEFA--GPWLIFSEISFISDVVN 352
 DB 307 EYPRWIRIPVNNRIAKVIRLFTGDSWLFISEVNFESNHN 350

RESULT 5

O60462
 ID O60462 PRELIMINARY; PRT; 931 AA.

AC O60462
 DT 01-AUG-1998 (TRENBLrel. 07, Created)

DT 01-AUG-1998 (TRENBLrel. 07, Last sequence update)
 DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)
 DE VASCULAR ENDOTHELIAL CELL GROWTH FACTOR 165 RECEPTOR 2.
 GN VEGF165R2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX TISSUE-BREAST;
 RX MEDLINE; 98188099.
 RA SOKER S., TAKASHIMA S., MIAO H.Q., NEUFELD G., KLAGSBRUN M.;
 RT "Neuropilin-1 is expressed by endothelial and tumor cells as an
 RL isoform-specific receptor for vascular endothelial growth factor.";
 RL Cell 92:735-745(1998).
 DR EMBL; AF016098; AAC12922.1; -;
 DR PFAM; PF00431; CUB; 2.
 DR PFAM; PF00754; F5_F8_type_C; 2.
 DR PFAM; PF00629; MAM; 1.
 SQ SEQUENCE 931 AA; 104830 MW; 912262F2 CRC32;

Query Match 10.2%; Score 222; DB 4; Length 931;
 Best Local Similarity 32.2%; Pred. No. 1.6e-11;
 Matches 65; Conservative 28; Mismatches 83; Indels 26; Gaps 8;

QY 7 HFDPA---KCRYALGMQDRTIPDSISASSWSDSI-AAHRSRLSSDGDGAWCPAGSVF 62
 DB 268 HOEPLNFQCNVPLGMEGRIANEQISASSTYSDGRWTPQSSRLHGD--DNGWTPN---L 322
 QY 63 PKEEYLQVLDLQRLHLVALVGTQGRHAGLGKE-----FSRSYRLVSRDGRWGMKDR 117
 DB 323 DSNKEYLVQDLRLTMTATATQ---GAISRETQYGVKSKLEVSTNGEDMWYRHG 378
 QY 118 WGOEIVISGNEDEPGVVKDLGPPVWRLVRYPRADRVMSVCLRVLYGLWRD----- 172
 DB 379 KNHKVFQANNDATEVVLNKLHAPLLTRFVRIRPOTWH-SGIALRLLEFGCRVTDAPCSNM 437
 QY 172 -GLLSYTPVPGQMYLSEAVYL 192
 DB 438 LGMLSLGLIADSQISASSTQEYL 459

RESULT 6

O14820
 ID O14820 PRELIMINARY; PRT; 909 AA.
 AC O14820;
 DT 01-JAN-1998 (TRENBLrel. 05, Created)
 DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
 DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)
 DE NEUROPILIN-2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 97470888.
 RA CHEN H., CHEDOTAL A., HE Z.-G., GOODMAN C.S., TESSIER-LAVIGNE M.;
 RT "Neuropilin-2, a novel member of the neuropilin family, is a high
 RT affinity receptor for the semaphorins Sema E and Sema IV but not Sema
 RT III.";
 RL Neuron 19:547-559(1997).
 DR EMBL; AF022859; AAC51788.1; -;
 DR PFAM; PF00431; CUB; 2.
 DR PFAM; PF00754; F5_F8_type_C; 2.
 DR PFAM; PF00629; MAM; 1.
 SQ SEQUENCE 909 AA; 102213 MW; A5F7D0CD CRC32;

Query Match 10.2%; Score 222; DB 4; Length 909;
 Best Local Similarity 32.2%; Pred. No. 1.6e-11;
 Matches 65; Conservative 28; Mismatches 83; Indels 26; Gaps 8;

```

QY 7 HFDPA---KCRYALGMQDRTIPDSISASSWSGST-AARHSRLSSDGDGAWCPAGSVF 62
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 268 HOEPLNFQCNVPLGMESGRIANEQISASTSYSDGRWTPQOSRLHGD--DNGWTPN---L 322

QY 63 PKEEYLOVDLQRLHLVALVGTQGRHAGLGKE-----FSRSYRLRYSDGRRWMWKDR 117
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 323 DSNKEYLOVDLRLTMTLTAITQ----GAISRETONGYVYKSYKLEVSTNGEDMMYVRHG 378

QY 118 WQVEISGNEDEPGVVLKDLGPPMVARLVRFYPRADRVMSVCLRVLYGCLWRD----- 172
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 379 KNHKVFOANNDATVVLNKLHAPLLTFRVIRPQTHW-SGIALRLLELFGCRVTDAPCSNM 437

QY 172 -GLLSYTPAPVGTWYLSAVYL 192
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 438 LGMLSGLIADSQISASTQEVYL 459

RESULT 7
ID O14821 PRELIMINARY; PRT; 926 AA.
AC O14821;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DE 01-MAY-1999 (TREMBlrel. 10, Last annotation update)
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RA CHEN H., CHEDOTAL A., HE Z.-G., GOODMAN C.S., TESSIER-LAVIGNE M.;
RT "Neuropilin-2, a novel member of the neuropilin family, is a high
RT affinity receptor for the semaphorins Sema E and Sema IV but not Sema
RL Neuron 19:547-559(1997).
DR EMBL; AF022854; AAC53377.1; -.
DR MGD; MGI:1100492; NRP2.
DR PFAM; PF00431; CUB; 2.
DR PFAM; PF00754; F5_F8_type_C; 2.
DR PFAM; PF00629; NAM; 1.
SQ SEQUENCE 926 AA; 104253 MW; 4C812073 CRC32;

```

```

Query Match 10.2%; Score 222; DB 4; Length 926;
Best Local Similarity 32.2%; Pred. No. 1.6e-11;
Matches 65; Conservative 28; Mismatches 83; Indels 26; Gaps 8;

QY 7 HFDPA---KCRYALGMQDRTIPDSISASSWSGST-AARHSRLSSDGDGAWCPAGSVF 62
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 268 HOEPLNFQCNVPLGMESGRIANEQISASTSYSDGRWTPQOSRLHGD--DNGWTPN---L 322

QY 63 PKEEYLOVDLQRLHLVALVGTQGRHAGLGKE-----FSRSYRLRYSDGRRWMWKDR 117
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 323 DSNKEYLOVDLRLTMTLTAITQ----GAISRETONGYVYKSYKLEVSTNGEDMMYVRHG 378

QY 118 WQVEISGNEDEPGVVLKDLGPPMVARLVRFYPRADRVMSVCLRVLYGCLWRD----- 172
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 379 KNHKVFOANNDATVVLNKLHAPLLTFRVIRPQTHW-SGIALRLLELFGCRVTDAPCSNM 437

QY 172 -GLLSYTPAPVGTWYLSAVYL 192
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 438 LGMLSGLIADSQISASTQEVYL 459

RESULT 8
ID O35373 PRELIMINARY; PRT; 909 AA.
AC O35373;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DE 01-MAY-1999 (TREMBlrel. 10, Last annotation update)
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RA CHEN H., CHEDOTAL A., HE Z.-G., GOODMAN C.S., TESSIER-LAVIGNE M.;
RT "Neuropilin-2, a novel member of the neuropilin family, is a high
RT affinity receptor for the semaphorins Sema E and Sema IV but not Sema
RL Neuron 19:547-559(1997).
DR EMBL; AF022855; AAC53378.1; -.
DR MGD; MGI:1100492; NRP2.
DR PFAM; PF00431; CUB; 2.
DR PFAM; PF00754; F5_F8_type_C; 2.
DR PFAM; PF00629; NAM; 1.
SQ SEQUENCE 909 AA; 103982 MW; 513B1B88 CRC32;

```

```

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RA CHEN H., CHEDOTAL A., HE Z.-G., GOODMAN C.S., TESSIER-LAVIGNE M.;
RT "Neuropilin-2, a novel member of the neuropilin family, is a high
RT affinity receptor for the semaphorins Sema E and Sema IV but not Sema
RL Neuron 19:547-559(1997).
DR EMBL; AF022854; AAC53377.1; -.
DR MGD; MGI:1100492; NRP2.
DR PFAM; PF00431; CUB; 2.
DR PFAM; PF00754; F5_F8_type_C; 2.
DR PFAM; PF00629; NAM; 1.
SQ SEQUENCE 909 AA; 102125 MW; CC2919A1 CRC32;

```

```

Query Match 10.1%; Score 220; DB 11; Length 909;
Best Local Similarity 31.9%; Pred. No. 2.4e-11;
Matches 59; Conservative 31; Mismatches 69; Indels 26; Gaps 8;

QY 7 HFDPA---KCRYALGMQDRTIPDSISASSWSGST-AARHSRLSSDGDGAWCPAGSVF 62
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 268 HOEPLNFQCNVPLGMESGRIANEQISASTSYSDGRWTPQOSRLHGD--DNGWTPN---L 322

QY 63 PKEEYLOVDLQRLHLVALVGTQGRHAGLGKE-----FSRSYRLRYSDGRRWMWKDR 117
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 323 DSNKEYLOVDLRLTMTLTAITQ----GAISRETONGYVYKSYKLEVSTNGEDMMYVRHG 378

QY 118 WQVEISGNEDEPGVVLKDLGPPMVARLVRFYPRADRVMSVCLRVLYGCLWRD----- 172
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 379 KNHKVFOANNDATVVLNKLHAPLLTFRVIRPQTHW-SGIALRLLELFGCRVTDAPCSNM 437

QY 172 -GLLS 175
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 438 LGMLS 442

```

```

RESULT 9
ID O35374 PRELIMINARY; PRT; 926 AA.
AC O35374;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DE 01-MAY-1999 (TREMBlrel. 10, Last annotation update)
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RA CHEN H., CHEDOTAL A., HE Z.-G., GOODMAN C.S., TESSIER-LAVIGNE M.;
RT "Neuropilin-2, a novel member of the neuropilin family, is a high
RT affinity receptor for the semaphorins Sema E and Sema IV but not Sema
RL Neuron 19:547-559(1997).
DR EMBL; AF022855; AAC53378.1; -.
DR MGD; MGI:1100492; NRP2.
DR PFAM; PF00431; CUB; 2.
DR PFAM; PF00754; F5_F8_type_C; 2.
DR PFAM; PF00629; NAM; 1.
SQ SEQUENCE 926 AA; 103982 MW; 513B1B88 CRC32;

```

```

Query Match 10.1%; Score 220; DB 11; Length 926;
Best Local Similarity 31.9%; Pred. No. 2.4e-11;
Matches 59; Conservative 31; Mismatches 69; Indels 26; Gaps 8;

```

```

Qy 7 HFDPA---KCRYALGMQDRTPDSDISASSWSGST-AARHSRLSSDGDGAWCPAGSVF 62.
D 268 HQEPPNFQCNVPLGMSGRIANEQISASSTFSDGRWTPQOSRLHGD--DNGWTPN---L 322
Qy 63 PKEEYLVQDLQRLHLVALVGTQGRHAGGLGKE-----FSRSYRLYSRDRRWGMWKDR 117
D 323 DSNKEYLVQDLRLFLTLTAITQ-----GAISRETQGYVYKSLKLEVTNGEDMVMYRHG 378
Qy 118 WGOEIVSGNEDPGVVLKDLGPPMVARLVFYPRADRVMSVCLRVLYGCLWRD----- 172
D 379 KNHIFQANDATEVVLNKLHMLPLLTFRIRPOTWH-LGIALRLFLGCRVTDAPCSNM 437
Qy 172 -GLLS 175
D 438 LGMLS 442

RESULT 10
O35375
AC O35375 PRELIMINARY; PRT; 931 AA.
DT 01-JAN-1998 (T-EMBLrel. 05, Created)
DT 01-JAN-1998 (T-EMBLrel. 05, Last sequence update)
DT 01-MAY-1999 (T-EMBLrel. 10, Last annotation update)
DE NEUROFILIN 2.
GN NRP2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=BALB/C;
RX MEDLINE; 97470888.
RA CHEN H., CHEDOTAL A., HE Z.-G., GOODMAN C.S., TESSIER-LAVIGNE M.;
RT "Neurofilin-2, a novel member of the neurofilin family, is a high
RT affinity receptor for the semaphorins Sema E and Sema IV but not Sema
RT III."
RL Neuron 19:547-559(1997).
DR EMBL; AF022857; AAC53380.1; -.
DR MGD; MGI:1100492; NRP2.
DR PFAM; PF00431; CUB; 2.
DR PFAM; PF00754; F5_F8_type_C; 2.
DR PFAM; PF00629; MAM; 1.
SQ SEQUENCE 931 AA; 104558 MW; 08BF6BBD CRC32;

Query Match 10.1%; Score 220; DB 11; Length 931;
Best Local Similarity 31.9%; Pred. No. 2.4e-11;
Matches 59; Conservative 31; Mismatches 69; Indels 26; Gaps 8;

Qy 7 HFDPA---KCRYALGMQDRTPDSDISASSWSGST-AARHSRLSSDGDGAWCPAGSVF 62
D 268 HQEPPNFQCNVPLGMSGRIANEQISASSTFSDGRWTPQOSRLHGD--DNGWTPN---L 322
Qy 63 PKEEYLVQDLQRLHLVALVGTQGRHAGGLGKE-----FSRSYRLYSRDRRWGMWKDR 117
D 323 DSNKEYLVQDLRLFLTLTAITQ-----GAISRETQGYVYKSLKLEVTNGEDMVMYRHG 378
Qy 118 WGOEIVSGNEDPGVVLKDLGPPMVARLVFYPRADRVMSVCLRVLYGCLWRD----- 172
D 379 KNHIFQANDATEVVLNKLHMLPLLTFRIRPOTWH-LGIALRLFLGCRVTDAPCSNM 437
Qy 172 -GLLS 175
D 438 LGMLS 442

RESULT 11
O35376
AC O35376 PRELIMINARY; PRT; 901 AA.
DT 01-JAN-1998 (T-EMBLrel. 05, Created)

```

```

DT 01-JAN-1998 (T-EMBLrel. 05, Last sequence update)
DT 01-MAY-1999 (T-EMBLrel. 10, Last annotation update)
DE NEUROFILIN 2.
GN NRP2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=BALB/C;
RX MEDLINE; 97470888.
RA CHEN H., CHEDOTAL A., HE Z.-G., GOODMAN C.S., TESSIER-LAVIGNE M.;
RT "Neurofilin-2, a novel member of the neurofilin family, is a high
RT affinity receptor for the semaphorins Sema E and Sema IV but not Sema
RT III."
RL Neuron 19:547-559(1997).
DR EMBL; AF022857; AAC53380.1; -.
DR MGD; MGI:1100492; NRP2.
DR PFAM; PF00431; CUB; 2.
DR PFAM; PF00754; F5_F8_type_C; 2.
DR PFAM; PF00629; MAM; 1.
SQ SEQUENCE 901 AA; 101402 MW; C8EB8DC3 CRC32;

Query Match 10.1%; Score 220; DB 11; Length 901;
Best Local Similarity 31.9%; Pred. No. 2.3e-11;
Matches 59; Conservative 31; Mismatches 69; Indels 26; Gaps 8;

Qy 7 HFDPA---KCRYALGMQDRTPDSDISASSWSGST-AARHSRLSSDGDGAWCPAGSVF 62
D 268 HQEPPNFQCNVPLGMSGRIANEQISASSTFSDGRWTPQOSRLHGD--DNGWTPN---L 322
Qy 63 PKEEYLVQDLQRLHLVALVGTQGRHAGGLGKE-----FSRSYRLYSRDRRWGMWKDR 117
D 323 DSNKEYLVQDLRLFLTLTAITQ-----GAISRETQGYVYKSLKLEVTNGEDMVMYRHG 378
Qy 118 WGOEIVSGNEDPGVVLKDLGPPMVARLVFYPRADRVMSVCLRVLYGCLWRD----- 172
D 379 KNHIFQANDATEVVLNKLHMLPLLTFRIRPOTWH-LGIALRLFLGCRVTDAPCSNM 437
Qy 172 -GLLS 175
D 438 LGMLS 442

RESULT 12
O35377
AC O35377 PRELIMINARY; PRT; 906 AA.
DT 01-JAN-1998 (T-EMBLrel. 05, Created)
DT 01-JAN-1998 (T-EMBLrel. 05, Last sequence update)
DT 01-MAY-1999 (T-EMBLrel. 10, Last annotation update)
DE NEUROFILIN 2.
GN NRP2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=BALB/C;
RX MEDLINE; 97470888.
RA CHEN H., CHEDOTAL A., HE Z.-G., GOODMAN C.S., TESSIER-LAVIGNE M.;
RT "Neurofilin-2, a novel member of the neurofilin family, is a high
RT affinity receptor for the semaphorins Sema E and Sema IV but not Sema
RT III."
RL Neuron 19:547-559(1997).
DR EMBL; AF022858; AAC53381.1; -.
DR MGD; MGI:1100492; NRP2.
DR PFAM; PF00431; CUB; 2.
DR PFAM; PF00754; F5_F8_type_C; 2.
DR PFAM; PF00629; MAM; 1.
SQ SEQUENCE 906 AA; 101978 MW; 5A716A03 CRC32;

```


Query Match 10.1%; Score 220; DB 11; Length 906;
 Best Local Similarity 31.9%; Pred. No. 2.4e-11;
 Matches 59; Conservative 31; Mismatches 69; Indels 26; Gaps 8;

QY 7 HFDPA---KCRYALGMDRTIPDSISASSWSST-AARHSRLSSDGDGAWCPAGSVF 62
 DB 268 HOEPENFOCNVPLMGESGRIANEOISASSTFSDGRWTPQOSRLHGD--DNGWTPN---L 322

QY 63 PREEYLVQDLQRLHLVALVGTQGRHAGLGKE-----FSRSYRLRYSRDGRWGWKDR 117
 DB 323 DSNKEYLVQDLFLMTLTAIAQ-----GAISRETQGYVVKSYKLEVTNGEDWVYRHG 378

QY 118 WQOEYVSGNDEPEGVLKDLGPPVAVLRVPRADRVMSVCLRVLYGCLWRD----- 172
 DB 379 KNHKIFQANNDATVVLNKLHMLPLTRFIRIPQWHL-LGIALRLFLGCRVTDAPCSNM 437

QY 172 -GLLS 175
 DB 438 LGMLS 442

RESULT 13
 O35378
 ID O35378 PRELIMINARY; PRT; 914 AA.
 AC O35378;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
 DE NEUROFILIN 2.
 GN NP2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/C;
 RX MEDLINE; 97470888.
 RA CHEN H., CHEDOTAL A., HE Z.-G., GOODMAN C.S., TESSIER-LAVIGNE M.;
 RT "Neurofilin-2, a novel member of the neurofilin family, is a high
 affinity receptor for the semaphorins Sema E and Sema IV but not Sema
 III.";
 RL Neuron 19:547-559(1997).
 DR EMBL; AF022861; AAC53382.1; -.
 DR MGD; MGI:1100492; NRP2.
 DR PFAM; PF00431; CUB; 2.
 DR PFAM; PF00754; F5_F8_type_C; 2.
 DR PFAM; PF00629; NAM; 1.
 SQ SEQUENCE 914 AA; 102701 MW; 5BE1E858 CRC32;

Query Match 10.1%; Score 220; DB 11; Length 914;
 Best Local Similarity 31.9%; Pred. No. 2.4e-11;
 Matches 59; Conservative 31; Mismatches 69; Indels 26; Gaps 8;

QY 7 HFDPA---KCRYALGMDRTIPDSISASSWSST-AARHSRLSSDGDGAWCPAGSVF 62
 DB 268 HOEPENFOCNVPLMGESGRIANEOISASSTFSDGRWTPQOSRLHGD--DNGWTPN---L 322

QY 63 PREEYLVQDLQRLHLVALVGTQGRHAGLGKE-----FSRSYRLRYSRDGRWGWKDR 117
 DB 323 DSNKEYLVQDLFLMTLTAIAQ-----GAISRETQGYVVKSYKLEVTNGEDWVYRHG 378

QY 118 WQOEYVSGNDEPEGVLKDLGPPVAVLRVPRADRVMSVCLRVLYGCLWRD----- 172
 DB 379 KNHKIFQANNDATVVLNKLHMLPLTRFIRIPQWHL-LGIALRLFLGCRVTDAPCSNM 437

QY 172 -GLLS 175
 DB 438 LGMLS 442

RESULT 14
 O35378
 ID O35378 PRELIMINARY; PRT; 914 AA.
 AC O35378;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
 DE NEUROFILIN 2.
 GN NP2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/C;
 RX MEDLINE; 97470888.
 RA CHEN H., CHEDOTAL A., HE Z.-G., GOODMAN C.S., TESSIER-LAVIGNE M.;
 RT "Neurofilin-2, a novel member of the neurofilin family, is a high
 affinity receptor for the semaphorins Sema E and Sema IV but not Sema
 III.";
 RL Neuron 19:547-559(1997).
 DR EMBL; AF022861; AAC53382.1; -.
 DR MGD; MGI:1100492; NRP2.
 DR PFAM; PF00431; CUB; 2.
 DR PFAM; PF00754; F5_F8_type_C; 2.
 DR PFAM; PF00629; NAM; 1.
 SQ SEQUENCE 914 AA; 102701 MW; 5BE1E858 CRC32;

Query Match 10.0%; Score 217; DB 11; Length 925;
 Best Local Similarity 31.9%; Pred. No. 4.5e-11;
 Matches 59; Conservative 31; Mismatches 69; Indels 26; Gaps 8;

QY 7 HFDPA---KCRYALGMDRTIPDSISASSWSST-AARHSRLSSDGDGAWCPAGSVF 62
 DB 268 HOEPENFOCNVPLMGESGRIANEOISASSTFSDGRWTPQOSRLHGD--DNGWTPN---V 322

QY 63 PREEYLVQDLQRLHLVALVGTQGRHAGLGKE-----FSRSYRLRYSRDGRWGWKDR 117
 DB 323 DSNKEYLVQDLFLMTLTAIAQ-----GAISRETQGYVVKSYKLEVTNGEDWVYRHG 378

QY 118 WQOEYVSGNDEPEGVLKDLGPPVAVLRVPRADRVMSVCLRVLYGCLWRD----- 172
 DB 379 KNHKIFQANNDATVVLNKLHMLPLTRFIRIPQWHL-LGIALRLFLGCRVTDAPCSNM 437

QY 172 -GLLS 175
 DB 438 LGMLS 442

RESULT 15
 O35276
 ID O35276 PRELIMINARY; PRT; 925 AA.
 AC O35276;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
 DE NEUROFILIN-2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SPRAGUE-DAWLEY;
 RX MEDLINE; 97433088.
 RA KOLODKIN A.L., LEVENGOOD D.V., ROWE E.G., TAI Y.-T., GIGER R.J.,
 RA GINTY D.D.;
 RT "Neurofilin is a semaphorin III receptor.";
 RL Cell 90:753-762(1997).
 DR EMBL; AF016297; AAC53338.1; -.
 DR PFAM; PF00431; CUB; 2.
 DR PFAM; PF00754; F5_F8_type_C; 2.
 DR PFAM; PF00629; NAM; 1.
 SQ SEQUENCE 925 AA; 103896 MW; B097CCCA CRC32;

Query Match 10.0%; Score 217; DB 11; Length 925;
 Best Local Similarity 31.9%; Pred. No. 4.5e-11;
 Matches 59; Conservative 31; Mismatches 69; Indels 26; Gaps 8;

QY 7 HFDPA---KCRYALGMDRTIPDSISASSWSST-AARHSRLSSDGDGAWCPAGSVF 62
 DB 268 HOEPENFOCNVPLMGESGRIANEOISASSTFSDGRWTPQOSRLHGD--DNGWTPN---V 322

QY 63 PREEYLVQDLQRLHLVALVGTQGRHAGLGKE-----FSRSYRLRYSRDGRWGWKDR 117
 DB 323 DSNKEYLVQDLFLMTLTAIAQ-----GAISRETQGYVVKSYKLEVTNGEDWVYRHG 378

QY 118 WQOEYVSGNDEPEGVLKDLGPPVAVLRVPRADRVMSVCLRVLYGCLWRD----- 172
 DB 379 KNHKIFQANNDATVVLNKLHMLPLTRFIRIPQWHL-LGIALRLFLGCRVTDAPCSNM 437

QY 172 -GLLS 175
 DB 438 LGMLS 442

RESULT 15
 O35276
 ID O35276 PRELIMINARY; PRT; 845 AA.
 AC O35276;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
 DE AEBPL.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-CANCELLOUS BONE;
 RX MEDLINE; 97079196.
 RA OHNO I., HASHIMOTO J., SHIMIZU K., TAKAKA K., OCHI T., MATSUBARA K.,
 RA OKUBO K.;
 RT "A cDNA cloning of human AEBPL from primary cultured osteoblasts and
 its expression in a differentiating osteoblastic cell line.";
 RL Biochem. Biophys. Res. Commun. 228:411-414(1996).
 DR EMBL; D86479; BAAL3094.1; -.
 DR PFAM; PF00754; F5_F8_type_C; 1.
 DR PFAM; PF00246; Zn_carboxypept; 2.
 SQ SEQUENCE 845 AA; 96173 MW; 81C8A087 CRC32;

Query Match 9.7%; Score 211.5; DB 4; Length 845;

Best Local Similarity 24.38; Pred. No. 1.2e-10;
Matches 80; Conservative 54; Mismatches 144; Indels 51; Gaps 12;

```
QY 9 DPACRYALGMQDRTIPDSIDISASSWSNSTAARHSLESSDG-----DGAWCPAGSV 61
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 68 EKVKCP-PIGMESHRIEDNQIRASSMLRHGLGAQRGLNMQTGATEDDDYIDGCAAEEDA 126
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 62 FPKEEYLQVDLQRHLVALVGTGGRHAGGLGKEFSRYRLYSRGRRWGWKDRWGOE 121
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 127 ---RTQWIEVDTRRTRTGTGVTGGRDS-SIHDDFVTFVFGESNDSQTWMTNGYEM 182
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 122 VISGEDPEGVVLKDLGPPWVARLVREYPRADRVMSVCLRVELYGCLWRDGLLSYAPVG 181
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 183 THGNVDKDTPVLSLPEPVVARIRIYPLTNW-GSLCMRLEVLCG-----SVAPVY 233
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 182 QTMYLSEAVYLNDSYDGTGGVQLYGGLGQLADGV-----YGLDDFRKSQEL 229
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 234 SYIAQNEVVATDDLDFRHS-----YKDMRQLMKVNEECPITIRYISLG-----KSSRGL 284
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 230 RWV-----PGYDVYVGSNHSFSSGVYEMEFEDRLRAEFAQMQVHCNNMHTLGARLPG 282
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 285 KIVAMEISDNPGCHELGEPEFRYTAGIHGNEV-LGRELLLLMLQYLCREYRGNFRVSL 343
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 283 V-ECFRFRGPAWAGEGEPMRHNLGNGLD 310
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 344 VODTRIHLVPSLNPDPYEEVAAQMGSEFGN 372
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

Search completed: November 4, 1999, 01:53:56
Job time: 7616 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 4, 1999, 01:54:22 ; Search time 15.86 Seconds
(without alignments)
711.163 Million cell updates/sec

Title: US-08-170-558-8
Perfect score: 2175
Sequence: 1 DADMKGHDPKCRYALGMQ.....LELEPRGQOPVAKPEGSPTA 399

Scoring table: BLOSUM62

Searched: 77977 seqs, 28268293 residues

Database : SwissProt_37:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2167	99.6	913	1	EDD1_HUMAN	Q08345 homo sapien
2	2021	92.9	910	1	EDD1_RAT	Q63474 rattus norv
3	1999.5	91.9	911	1	EDD1_MOUSE	Q03146 mus musculu
4	1002.5	46.1	855	1	TRK3_HUMAN	Q16832 homo sapien
5	995.5	45.8	854	1	TRK3_MOUSE	Q62371 mus musculu
6	213	9.8	928	1	NRX_XENLA	P28824 xenopus lae
7	210	9.7	2133	1	FA8_PIG	P12263 sus scrofa
8	204	9.4	923	1	NRX_MOUSE	P97333 mus musculu
9	192	8.8	2351	1	FA8_HUMAN	P00451 homo sapien
10	189	8.7	914	1	NRX_CHICK	P79795 gallus gall
11	178.5	8.2	2224	1	FA5_HUMAN	P12259 homo sapien
12	177.5	8.2	387	1	MFGM_HUMAN	Q08431 homo sapien
13	175	8.0	2319	1	FA8_MOUSE	Q06194 mus musculu
14	174.5	8.0	2211	1	FA5_BOVIN	Q28107 bos taurus
15	174.5	8.0	463	1	MFGM_MOUSE	P21956 mus musculu
16	172.5	7.9	427	1	MFGM_RAT	P70490 rattus norv
17	167.5	7.7	427	1	MFGM_BOVIN	Q95114 bos taurus
18	166.5	7.7	409	1	MFGM_PIG	P79385 sus scrofa
19	145.5	6.7	3133	1	HMCT_BOMMO	P98092 bombyx mori
20	143	6.6	257	1	DIS2_DICDI	P42530 dictyosteli
21	106.5	4.9	446	1	HGD_LEMNI	Q00667 emericella
22	97	4.5	1101	1	FURL_DROME	P30430 drosophila
23	97	4.5	899	1	FURL_DROME	P26016 drosophila
24	97	4.5	789	1	LP1B_DROME	P11996 drosophila
25	96	4.4	253	1	DISC_DICDI	P02887 dictyosteli
26	92.5	4.3	985	1	AGLU_ASPNG	P56526 aspergillus
27	92	4.2	635	1	SIR_SCHY3	P72854 synchocyst
28	91.5	4.2	411	1	CYB_HNRV1	Q31215 chromatiu
29	91	4.2	253	1	DISA_DICDI	P02886 dictyosteli
30	89	4.1	477	1	CSCA_ECOLI	P40714 escherichia
31	89	4.1	867	1	SSPO_BOVIN	P98167 bos taurus
32	89	4.1	3164	1	TEGU_HSV11	P10220 herpes simp
33	87.5	4.0	139	1	LAMP_PETMA	P33575 petromyzon
34	87.5	4.0	1168	1	MYSC_ACACA	P10569 acanthamoeb
35	87	4.0	439	1	HFH1_RAT	Q63244 rattus norv
36	86.5	4.0	149	1	DISD_DICDI	P02888 dictyosteli
37	86.5	4.0	379	1	UL28_HCMVA	P16847 human cytom
38	86	4.0	253	1	RM02_ACACA	P46763 acanthamoeb
39	85.5	3.9	3567	1	ERY2_SACER	Q03132 saccharomyc
40	85.5	3.9	1460	1	RPC1_YEAST	P04051 saccharomyc
41	84.5	3.9	393	1	CYB_METSE	Q47499 metridium s
42	84	3.9	548	1	FLR1_YEAST	P38124 saccharomyc
43	84	3.9	416	1	HXD3_HUMAN	P31249 homo sapien

ALIGNMENTS

```
RESULT 1
ID EDD1_HUMAN STANDARD; PRT: 913 AA.
AC Q08345; Q14196; Q16562;
DT 01-FEB-1995 (REL. 31, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE EPIHELIAL DISCOIDIN DOMAIN RECEPTOR 1 PRECURSOR (EC 2.7.1.112)
DE (TYROSINE-PROTEIN KINASE CAK) (CELL ADHESION KINASE) (TYROSINE KINASE
DE (DISCOIDIN RECEPTOR TYROSINE KINASE) (TRK E) (PROTEIN-TYROSINE
DE KINASE RTK 6).
GN EDD1 OR CAK OR DDR1 OR TRKE OR RTK6.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=FETAL LIVER;
RX MEDLINE; 94134417.
RA PEREZ J.L., SHEN X., FINKERNAGEL S., SCIORRA L., JENKINS N.A.,
RA GILBERT D.J., COELAND N.G., WONG T.W.;
RT "Identification and chromosomal mapping of a receptor tyrosine kinase
RT with a putative phospholipid binding sequence in its ectodomain.";
RN ONCOGENE 9:211-219(1994).
RP SEQUENCE FROM N.A.
RX MEDLINE; 93296201.
RA JOHNSON J.D., EDMAN J.C., RUTTER W.J.;
RT "A receptor tyrosine kinase found in breast carcinoma cells has an
RT extracellular discoidin I-like domain.";
RN PROC. NATL. ACAD. SCI. U.S.A. 90:5677-5681(1993).
RP SEQUENCE FROM N.A.
RX MEDLINE; 97131588.
RA SAKUMA S., TADA M., SAYA H., SAWAMURA Y., SHINOHE Y., ABE H.;
RT "Receptor protein tyrosine kinase DDR is up-regulated by p53
RT protein.";
RN FEBS LETT. 398:165-169(1996).
RP SEQUENCE FROM N.A.
RX MEDLINE; 96389017.
RA PLAYFORD M.P., BUTLER R.J., WANG X.C., KATSO R.M., COOKE I.E.,
RA GANESAN T.S.;
RT "The genomic structure of discoidin receptor tyrosine kinase.";
RN GENOME RES. 6:620-627(1996).
RP SEQUENCE FROM N.A. (SHORT FORM).
RC TISSUE=LUNG;
RX MEDLINE; 96204002.
RA PEREZ J.L., JING S.O., WONG T.W.;
RT "Identification of two isoforms of the Cak receptor kinase that are
RT coexpressed in breast tumor cell lines.";
RN ONCOGENE 12:1469-1477(1996).
RP SEQUENCE FROM N.A. (SHORT FORM).
RC TISSUE=BRAIN, AND KERATINOCYTES;
RX MEDLINE; 94043265.
RA DI MARCO E., CUTULI N., GUERRA L., CANCEDDA R., DE LUCA M.;
RT "Molecular cloning of trke, a novel trk-related putative tyrosine
RT kinase receptor isolated from normal human keratinocytes and widely
RT expressed by normal human tissues.";
RN J. BIOL. CHEM. 268:24290-24295(1993).
RP SEQUENCE FROM N.A. (SHORT FORM).
RC TISSUE-OVARY;
```

RX MEDLINE: 95151638.
 RA LAVAL S., BUTLER R., SHELLING A.N., HANBY A.M., POULSOM R.,
 GNANESAN T.S.,
 RT "Isolation and characterization of an epithelial-specific receptor
 tyrosine kinase from an ovarian cancer cell line.";
 RL CELL GROWTH DIFFER. 5:1173-1183(1994).
 CC -!- FUNCTION: MAY BE INVOLVED IN CELL-CELL INTERACTIONS AND
 RECOGNITION.
 CC -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE - ADP +
 PROTEIN TYROSINE PHOSPHATE.
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -!- TISSUE SPECIFICITY: EXPRESSED AT LOW LEVELS IN MOST ADULT
 TISSUES AND IS HIGHEST IN THE BRAIN AND LUNG. ABUNDANT IN
 BREAST CARCINOMA CELL LINES.
 CC -!- DOMAIN: THE GLY/PRO-RICH DOMAINS MAY BE REQUIRED FOR AN UNUSUAL
 GEOMETRY OF INTERACTION WITH LIGAND OR SUBSTRATES.
 CC -!- ALTERNATIVE PRODUCTS: TWO ISOFORMS ARE PRODUCED BY ALTERNATIVE
 SPLICING. THE LONG FORM (CALLED CAK I AND SHOWN HERE) AND THE
 SHORT FORM (CALLED CAK II) ONLY DIFFER BY THE ABSENCE OF A 37
 RESIDUES SEGMENT.
 CC -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
 DOMAIN.
 CC -!- SIMILARITY: CONTAINS 1 F5/8 TYPE C DOMAIN.
 CC -!- SIMILARITY: BELONGS TO THE INSULIN RECEPTOR FAMILY OF TYROSINE-
 PROTEIN KINASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L11315; G403387; -
 DR EMBL: L20817; G306475; -
 DR EMBL: U48705; G1754536; -
 DR EMBL: X98208; E251231; -
 DR EMBL: X99023; E251231; JOINED.
 DR EMBL: X99024; E251231; JOINED.
 DR EMBL: X99025; E251231; JOINED.
 DR EMBL: X99026; E251231; JOINED.
 DR EMBL: X99027; E251231; JOINED.
 DR EMBL: X99028; E251231; JOINED.
 DR EMBL: X99029; E251231; JOINED.
 DR EMBL: X99030; E251231; JOINED.
 DR EMBL: X99031; E251231; JOINED.
 DR EMBL: X99032; E251231; JOINED.
 DR EMBL: X99033; E251231; JOINED.
 DR EMBL: X99034; E251231; JOINED.
 DR EMBL: L57508; G1160925; -
 DR EMBL: X74979; G400463; -
 DR EMBL: Z29093; G732800; -
 DR MIM: 600432; -
 DR MIM: 600408; -
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00239; RECEPTOR_TYR_KIN_II; 1.
 DR PROSITE: PS01285; FA58C_1; 1.
 DR PROSITE: PS01286; FA58C_2; 1.
 DR PFAM: PF00069; pkinase; 1.
 DR PFAM: PF00754; F5_F8_type_C; 1.
 DR HSSP: P11362; 1FGI.
 KW TRANSFERASE: TYROSINE-PROTEIN KINASE; GLYCOPROTEIN; SIGNAL;
 KW PHOSPHORYLATION; TRANSMEMBRANE; RECEPTOR; ATP-BINDING;
 KW ALTERNATIVE SPLICING.
 FT SIGNAL 1 18 POTENTIAL.
 FT CHAIN 19 913 EPITHELIAL DISCOIDIN DOMAIN RECEPTOR 1.
 FT DOMAIN 19 416 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 417 443 POTENTIAL.
 FT DOMAIN 444 913 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 31 185 F5/8 TYPE C (PHOSPHOLIPID-BINDING,

FT DOMAIN 377 415 POTENTIAL).
 FT DOMAIN 476 601 GLY/PRO-RICH.
 FT DOMAIN 610 905 GLY/PRO-RICH.
 FT NP_BIND 616 624 PROTEIN KINASE.
 FT BINDING 655 655 ATP (BY SIMILARITY).
 FT ACT_SITE 766 766 ATP (BY SIMILARITY).
 FT DISULFID 31 185 BY SIMILARITY.
 FT MOD_RES 513 513 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT MOD_RES 792 792 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT MOD_RES 796 796 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT MOD_RES 797 797 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT CARBOHYD 211 211 POTENTIAL.
 FT CARBOHYD 260 260 POTENTIAL.
 FT CARBOHYD 371 371 POTENTIAL.
 FT CARBOHYD 394 394 POTENTIAL.
 FT CARSPLIC 506 542 MISSING (IN SHORT FORM).
 FT CONFLICT 94 94 L -> V (IN REF. 2 AND 3).
 FT CONFLICT 833 833 L -> V (IN REF. 2 AND 3).
 FT CONFLICT 847 867 QLTDEQVIVNAGEFFRDGGRQ -> SAHRRAGHRERGVL
 GPQA (IN REF. 4).
 FT SEQUENCE 913 AA: 101127 MW: 16071364 CRC32:
 Query Match 99.6%; Score 2167; DB 1; Length 913;
 Best Local Similarity 99.7%; Pred. No. 1.1e-170;
 Matches 398; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 DADMGHFDPAKCRVAGLGMODRTIPDSISASSWSNSTAAHRSLESDDGAGWCPAGS 60
 DB 19 DADMGHFDPAKCRVAGLGMODRTIPDSISASSWSNSTAAHRSLESDDGAGWCPAGS 78
 QY 61 VFPKEEYLVQDLQRLHLVALVGTQGRHAGGLGKEFSYRLRYSDGRRMMGWKDRWGQ 120
 DB 79 VFPKEEYLVQDLQRLHLVALVGTQGRHAGGLGKEFSYRLRYSDGRRMMGWKDRWGQ 138
 QY 121 EVISGNEDEPVGVLKDLGPPMVARLVRFPYPRADRVMSVCLRVLYGCLWRDGLLSYAPV 180
 DB 139 EVISGNEDEPVGVLKDLGPPMVARLVRFPYPRADRVMSVCLRVLYGCLWRDGLLSYAPV 198
 QY 181 GQTMYLEAVYLVNDSTYDGHVGGYGLGQGLADGVVGLDDFRKLSQELRVWPGYDYVGW 240
 DB 199 GQTMYLEAVYLVNDSTYDGHVGGYGLGQGLADGVVGLDDFRKLSQELRVWPGYDYVGW 258
 QY 241 SNHSFSGYVMEFEFDRLEAFQAMOVHCHNMHTLGARLPGGVECFRRGRPAMAVEGEPM 300
 DB 259 SNHSFSGYVMEFEFDRLEAFQAMOVHCHNMHTLGARLPGGVECFRRGRPAMAVEGEPM 318
 QY 301 RHNLGNLGDPRARAVSVPLGGRVAFRLQCLRFAGPWLFFSEISFISDVVNNSSPALGG 360
 DB 319 RHNLGNLGDPRARAVSVPLGGRVAFRLQCLRFAGPWLFFSEISFISDVVNNSSPALGG 378
 QY 361 TFPAPWPPGPPPTNFSSLELEPRGQQPVAKPGESPTA 399
 DB 379 TFPAPWPPGPPPTNFSSLELEPRGQQPVAKPGESPTA 417
 RESULT 2
 EDDL_RAT
 ID EDDL_RAT STANDARD; PRT; 910 AA.
 AC 063474;
 DT 01-NOV-1997 (REL. 35, CREATED)
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
 DE EPITHELIAL DISCOIDIN DOMAIN RECEPTOR 1 (PRECURSOR (EC 2.7.1.112)
 DE (TYROSINE-PROTEIN KINASE CAK) (CELL ADHESION KINASE) (TYROSINE KINASE
 DE DDR) (DISCOIDIN RECEPTOR TYROSINE KINASE) (PROTEIN-TYROSINE KINASE
 DE PK-3).
 GN EDDL OR PK3.
 OS RATTUS NORVEGICUS (RAT).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.
 RN [1]

SEQUENCE FROM N.A.
 RC STRAIN-SPRAGUE-DAWLEY; TISSUE=BRAIN;
 RA MEDLINE; 94173920.
 RA SANCHEZ M.P., TAPLEY P., SAINI S.S., HE B., PULIDO D., BARBACID M.;
 RT "Multiple tyrosine protein kinases in rat hippocampal neurons:
 RT isolation of PK-3, a receptor expressed in proliferative zones of
 RT the developing brain";
 RL PROC. NATL. ACAD. SCI. U.S.A. 91:1819-1823(1994).
 CC -1- FUNCTION: MAY BE INVOLVED IN CELL-CELL INTERACTIONS AND
 CC RECOGNITION (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
 CC PROTEIN TYROSINE PHOSPHATE.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: VARIOUS EMBRYONIC AND ADULT TISSUES; ALSO
 CC PROLIFERATIVE ZONES OF THE DEVELOPING BRAIN; HIPPOCAMPAL NEURONS.
 CC -1- SIMILARITY: WITH THE CONSERVED CATALYTIC DOMAINS OF TYR-PROTEIN
 CC KINASES.
 CC -1- SIMILARITY: CONTAINS 1 F5/8 TYPE C DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE INSULIN RECEPTOR FAMILY OF TYROSINE-
 CC PROTEIN KINASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L26525; G432481;
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.
 DR PROSITE; PS01285; FA58C_1; 1.
 DR PROSITE; PS01286; FA58C_2; 1.
 DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
 DR PFAM; PF00069; pkinase; 1.
 DR PFAM; PF00754; F5_F8_type_C; 1.
 DR HSP; P11362; 1FG1.
 KW TRANSFERASE; TYROSINE-PROTEIN KINASE; GLYCOPROTEIN; SIGNAL;
 KW PHOSPHORYLATION; TRANSMEMBRANE; RECEPTOR; ATP-BINDING.
 FT SIGNAL 1 19
 FT CHAIN 20 910
 FT DOMAIN 20 413
 FT TRANSMEM 414 440
 FT DOMAIN 441 910
 FT DOMAIN 32 186
 FT DOMAIN 378 412
 FT DOMAIN 473 598
 FT DOMAIN 607 902
 FT NP_BIND 613 621
 FT BINDING 652 652
 FT ACT_SITE 763 763
 FT DISULFID 32 186
 FT MOD_RES 510 510
 FT MOD_RES 789 789
 FT MOD_RES 793 793
 FT MOD_RES 794 794
 FT CARBOHYD 212 212
 FT CARBOHYD 261 261
 FT CARBOHYD 371 371
 FT CARBOHYD 391 391
 SQ SEQUENCE 910 AA; 101164 MW; D69094A9 CRC32;
 Query Match 92.9%; Score 2021; DB 1; Length 910;
 Best Local Similarity 93.0%; Pred. No. 1.2e-158;
 Matches 371; Conservative 9; Mismatches 15; Indels 4; Gaps 1;
 QY 1 DADMKGHFDPAKRYALGMQDRTPDSDISASSWSWSDTAARHSRLSSDGDGAWCPAGS 60
 DB 20 DADMKGHFDPAKRYALGMQDRTPDSDISVSSWSWSDTAARHSRLSSDGDGAWCPAGF 79

QY 61 VPKKEEYLOVDLQRLHLVALVGTQGRHAGGLGKEFSYRLRYSDGRRWMGWKDRWGQ 120
 DB 80 VPKKEEYLOVDLQRLHLVALVGTQGRHAGGLGKEFSYRLRYSDGRRWMGWKDRWGQ 139
 QY 121 EVISGNEDEPGVVLKDLGPPMVARLVREFYPRADRVMSVCLRVLYGCLWRDGLLSYTA 180
 DB 140 EVISGNEDEPGVVLKDLGPPMVARLVREFYPRADRVMSVCLRVLYGCLWRDGLLSYTA 199
 QY 181 GQTMWSEAYVINDSTYDGHVGGLOGLGOLADGVVGLDDFRKSOELRWPGYDVGW 240
 DB 200 GQTMOLSEMYVINDSTYDGHVGGLOGLGOLADGVVGLDDFRKSOELRWPGYDVGW 259
 QY 241 SNHSFSSGVEVEFEEDRLRAFOAMOVHCNNMHTLGARLPGVVECRFRGPPAMAWEGEP 300
 DB 260 SNHSFSSGVEVEFEEDRLRSFQTMVHCNNMHTLGARLPGVVECRFRGPPAMAWEGEP 319
 QY 301 RNLGNLGDPRARAVSVPLGGRVAREFLOCRFLFAGPWLFLSEISFISDVVNSSPALGG 360
 DB 320 HHALGSLGDPRARAVSVPLGGRVAREFLOCRFLFAGPWLFLSEISFISDVVNSSPALGG 375
 QY 361 TFPAPMPPGPPPTNFSSLELEPRGQOPVAKPEGSP 399
 DB 376 TFPAPMPPGPPPTNFSSLELEPRGQOPVAKPEGSP 414
 RESULT 3
 EDDL_MOUSE
 ID EDDL_MOUSE STANDARD; PRT; 911 AA.
 AC Q03146;
 DT 01-OCT-1994 (REL. 30, CREATED)
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE EPIHELIAL DISCOIDIN DOMAIN RECEPTOR 1 PRECURSOR (EC 2.7.1.112)
 DE (TYROSINE-PROTEIN KINASE CAK) (CELL ADHESION KINASE) (TYROSINE KINASE
 DE DDR) (DISCOIDIN RECEPTOR TYROSINE KINASE) (PROTEIN-TYROSINE KINASE
 DE MPK-6).
 GN EDDR1 OR CAK OR MPK6.
 OS MUS MUSCULUS (MOUSE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MORINAE; MUS.
 [1]
 SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
 RN STRAIN-C57BL;
 RC MEDLINE; 96204002.
 RA PEREZ J.L., JING S.Q., WONG T.W.;
 RT "Identification of two isoforms of the Cak receptor kinase that are
 RT coexpressed in breast tumor cell lines.";
 RL ONCOGENE 12:1469-1477(1996).
 [2]
 SEQUENCE OF 766-822 FROM N.A.
 RC STRAIN-C57BL; TISSUE=EMBRYONIC BRAIN;
 RX MEDLINE; 93096484.
 RA GILARDI-HEBENSTREIT P., NIETO M.A., FRAIN M., MATTEI M.-G.,
 RA CHESTER A., WILKINSON D.G., CHARNAY P.;
 RT "An Eph-related receptor protein tyrosine kinase gene segmentally
 RT expressed in the developing mouse hindbrain.";
 RL ONCOGENE 7:2499-2506(1992).
 CC -1- FUNCTION: MAY BE INVOLVED IN CELL-CELL INTERACTIONS AND
 CC RECOGNITION (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
 CC PROTEIN TYROSINE PHOSPHATE.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: THE PREDOMINANT ISOFORM CAK I, IS EXPRESSED IN
 CC DEVELOPING EMBRYO AND ADULT BRAIN; CAK II IS EXPRESSED IN VARIOUS
 CC EPITHELIAL CELLS.
 CC -1- ALTERNATIVE PRODUCTS: TWO ISOFORMS ARE PRODUCED BY ALTERNATIVE
 CC SPLICING. THE LONG FORM (CALLED CAK I AND SHOWN HERE) AND THE
 CC SHORT FORM (CALLED CAK II) ONLY DIFFER BY THE ABSENCE OF A 37
 CC RESIDUES SEGMENT.
 CC -1- SIMILARITY: WITH THE CONSERVED CATALYTIC DOMAINS OF TYR-PROTEIN
 CC KINASES.
 CC -1- SIMILARITY: CONTAINS 1 F5/8 TYPE C DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE INSULIN RECEPTOR FAMILY OF TYROSINE-

1	Qy	DADMGKGFDP	PAKRCYALCMQDRTIP	DSISASSWSNSTAARHSRLSSDGDCAWCP	PAGS	60
20	Db	DADMGKGFDP	PAKRCYALCMQDRTIP	DSISVSSWSNSTAARHSRLSSDGDCAWCP	PAGP	79
61	Qy	VFPKEEYLVQ	DLRHLHVALVGTQGR	HAGGLGKEFSRSYRLRSRDRRWMGKDRWGQ		120
80	Db	VFPKEEYLVQ	DLRHLHVALVGTQGR	HAGGLGKEFSRSYRLRSRDRRWMGKDRWGQ		139
121	Qy	EVISGNEDEGV	VVLKDLGPPMVARLVRY	PRADRVMSVCLRVELYGCLRDRGLLSYTAPV		180
140	Db	EVISGNEDEGV	VVLKDLGPPMVARLVRY	PRADRVMSVCLRVELYGCLRDRGLLSYTAPV		199
181	Qy	GOTMYLSEA-	VTYLNSTYDGTGGVGG	LQGLADGVVGLDDPRKSOELRVWPGYDVG		239
200	Db	GOTMYLSEA-	VTYLNSTYDGTGGVGG	LQGLADGVVGLDDPRKSOELRVWPGYDVG		259
240	Qy	WSNHSPSSCY	VMSEFEEDRLRAFOAMOV	HCNMHMTGLARLPGGVCECFRRGPAMAWBCEP		299

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation in the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/announcement/](http://www.isb-sib.ch/announcement/or send an email to license@isb-sib.ch) or send an email to license@isb-sib.ch).

CC	EMBL; X74764; G433338; -
DR	MIM; 191311 -
DR	PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR	PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.
DR	PROSITE; PS01285; FA58C_1; 1.
DR	PROSITE; PS01286; FA58C_2; 1.
DR	PROSITE; PS50011; PROTEIN_KINASE_DOWN; 1.
DR	PFAM; PF00069; pkinase; 1.
DR	PFAM; PF00754; F5_P8_type_C; 1.
DR	HSPF; P11362; IRTG1.
DR	TRANSFERASE; TYROSINE-PROTEIN KINASE; GLYCOPROTEIN; SIGNAL;
KW	PHOSPHORYLATION; TRANSMEMBRANE; RECEPTOR; ATP-BINDING.
FT	SIGNAL 1 21
FT	CHAIN 22 855
FT	DOMAIN 22 399
FT	RECEPTOR PROTEIN-TYROSINE KINASE TKT.
FT	EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 400 421 POTENTIAL.
 FT DOMAIN 422 855 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 30 185 F5/8 TYPE C.
 FT DOMAIN 563 ? PROTEIN KINASE.
 FT NP_BIND 569 577 ATP (BY SIMILARITY).
 FT BINDING 608 608 ATP (BY SIMILARITY).
 FT ACT_SITE 710 710 BY SIMILARITY.
 FT DISULFID 30 185 BY SIMILARITY.
 FT MOD_RES 740 740 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT CARBOHYD 121 121 POTENTIAL.
 FT CARBOHYD 213 213 POTENTIAL.
 FT CARBOHYD 261 261 POTENTIAL.
 FT CARBOHYD 280 280 POTENTIAL.
 FT CARBOHYD 372 372 POTENTIAL.
 SQ SEQUENCE 855 AA; 96752 MW; 2623D841 CRC32;

Query Match 46.1%; Score 1002.5; DB 1; Length 855;
 Best Local Similarity 53.6%; Pred. No. 6.4e-75;

Matches 196; Conservative 54; Mismatches 109; Indels 7; Gaps 6;

QY 5 KGHFDPKCRYALGMODRTIPDSISASSWSDDTAARHSRLSSDGDGAWCPAGSVFPK 64
 Db 22 KAQVNPACRYPLGSGGGIPDEDITASSOWSESTAAYKGRDSEGDGAWCPPIPVDP 81
 QY 65 E-EYLVQDQLRLHLVALVGTGRHAGGLGKEFSRSLRYSDRGRWMGKDRWGQEV 123
 Db 82 DLKEFLQIDLRLHLVLTITLVTGTRHAGGIEFAPMYKINYSRDRGRWMGKDRWGQEV 141
 QY 124 SGNDEPEGVVLKDLGPPMVARLVRYPRADRVMSCLVRLVGLCWRDGLLSYAPVQOT 183
 Db 142 DGNPNFYDIFLADLEPPIVARVRIPYVDHSMNVMCRVELYGCWLDGLVSNAPAGQ 201
 QY 184 MYL--SEAVYLDNDYDHTVGLGYQLGQADGVVGLDDFRKSOELRVWPGYDYGWS 241
 Db 202 FVLPGGSIYLDNDYD--AVGYSMTGELGQLTDCVSGLDQFTQHEVHVWPGYDYGWR 260
 QY 242 NSEFSGVYEMEFEDRLRAQAMQVHNNHTLGLARLPGGVECFRFRGPPAMWEGEPMR 301
 Db 261 NESATNGIEINFEFDRINFTMKVHNNMFAKGVKIFKEVQCYF-RSEASEWEPNAIS 319
 QY 302 HNIGNLGDPRARVSVPLGGVRVFLQCRFLFAGPWLLEIFSEIFSD-VYNNSSPALGG 360
 Db 320 FPLVLDNDVNSPASFVTVPLHHEMSAIAKQHFADTWMFSEITFQSDAAMINSEAL-P 378
 QY 361 FPPAP 366
 Db 379 TSPMAP 384

RESULT 5
 TRK3_MOUSE
 ID TRK3_MOUSE STANDARD; PRT; 854 AA.
 AC Q62371;
 DT 01-NOV-1997 (REL. 35, CREATED)
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE RECEPTOR PROTEIN-TYROSINE KINASE TKT PRECURSOR (EC 2.7.1.112)
 DE (TYROSINE-PROTEIN KINASE TYRO 10) (NEUROTROPHIC TYROSINE KINASE,
 DE RECEPTOR-RELATED 3).
 GN NTRK3 OR TKT OR TYRO10.
 OS MUS MUSCULUS (MOUSE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 94057796.
 RA KARN T., HOLTRICH U., BRAEUNINGER A., BOEHME B., WOLF G.,
 RA RUEBSAMEN-WAIGMANN H., STREIBHARDT K.;
 RT "Structure, expression and chromosomal mapping of TKT from man and
 RT mouse: a new subclass of receptor tyrosine kinases with a factor
 RT VIII-like domain."
 RL ONCOGENE 8:3433-3440(1993).

RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-C57BL/6; TISSUE-BRAIN;
 RX MEDLINE; 94151011.
 RA LAI C., LEMKE G.E.;
 RT "Structure and expression of the Tyro 10 receptor tyrosine kinase."
 RL ONCOGENE 9:877-883(1994).
 CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE -
 CC ADP + PROTEIN TYROSINE PHOSPHATE.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED; HIGH LEVELS IN SKELETAL
 CC MUSCLE, HEART, CNS, AND KIDNEY; LESS IN OTHER TISSUES. THE MAJOR
 CC 10 KD TRANSCRIPT IS EXPRESSED IN HIGH LEVELS IN HEART AND LUNG,
 CC LESS IN BRAIN AND TESTIS.
 CC -1- ALTERNATIVE PRODUCTS: DIFFERENT TRANSCRIPTS ARE DERIVED FROM ONE
 CC GENE.
 CC -1- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
 CC DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 F5/8 TYPE C DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE INSULIN RECEPTOR FAMILY OF TYROSINE-
 CC PROTEIN KINASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X76505; G435162; ALT_INIT.
 CC MGD; MGI:97385; NTRK3.
 CC DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 CC DR PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.
 CC DR PROSITE; PS01285; FA58C_1; 1.
 CC DR PROSITE; PS01286; FA58C_2; 1.
 CC DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 CC DR PFAM; PF00069; pkinase; 1.
 CC DR PFAM; PF00754; F5_F8_type_C; 1.
 CC DR HSP; P11362; IFGI.
 CC TRANSFERASE; TYROSINE-PROTEIN KINASE; GLYCOPROTEIN; SIGNAL;
 CC PHOSPHORYLATION; TRANSMEMBRANE; RECEPTOR; ATP-BINDING.
 KW SIGNAL 1 21
 FT CHAIN 22 854 RECEPTOR PROTEIN-TYROSINE KINASE TKT.
 FT DOMAIN 22 399 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 400 421 POTENTIAL.
 FT DOMAIN 422 854 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 30 185 F5/8 TYPE C.
 FT DOMAIN 563 ? PROTEIN KINASE.
 FT NP_BIND 569 577 ATP (BY SIMILARITY).
 FT BINDING 608 608 ATP (BY SIMILARITY).
 FT ACT_SITE 709 709 BY SIMILARITY.
 FT DISULFID 30 185 BY SIMILARITY.
 FT MOD_RES 739 739 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT CARBOHYD 121 121 POTENTIAL.
 FT CARBOHYD 213 213 POTENTIAL.
 FT CARBOHYD 261 261 POTENTIAL.
 FT CARBOHYD 280 280 POTENTIAL.
 FT CARBOHYD 372 372 POTENTIAL.
 SQ SEQUENCE 854 AA; 96482 MW; B78B6551 CRC32;

Query Match

Best Local Similarity 45.8%; Score 995.5; DB 1; Length 854;

Matches 198; Conservative 58; Mismatches 116; Indels 27; Gaps 7;

QY 5 KGHFDPKCRYALGMODRTIPDSISASSWSDDTAARHSRLSSDGDGAWCPAGSVFPK 64
 Db 22 KAQVNPACRYPLGSGGGIPDEDITASSOWSESTAAYKGRDSEGDGAWCPPIPVDP 81
 QY 65 E-EYLVQDQLRLHLVALVGTGRHAGGLGKEFSRSLRYSDRGRWMGKDRWGQEV 123
 Db 82 DLKEFLQIDLRLHLVLTITLVTGTRHAGGIEFAPMYKINYSRDRGRWMGKDRWGQEV 141

RA YOUSOUFIAN H., ANTONARAKIS S.E., BELL W., GRIFFIN A.M.,
 RT KAZAZIAN H.H.;
 RA "Nonsense and missense mutations in hemophilia A: estimate of the
 RT relative mutation rate at CG dinucleotides";
 RL AM. J. HUM. GENET. 42:718-725(1988).
 RN [15]
 RP VARIANT GLY-291.
 RX MEDLINE; 88220354.
 RA YOUSOUFIAN H., WONG C., ARONIS S., PLATOKOUKIS H., KAZAZIAN H.H. JR.,
 RA ANTONARAKIS S.E.;
 RT "Moderately severe hemophilia A resulting from Glu-->Gly substitution
 RT in exon 7 of the factor VIII gene";
 RL AM. J. HUM. GENET. 42:867-871(1988).
 RN [16]
 RP VARIANT CYS-1708.
 RX MEDLINE; 89274393.
 RA O'BRIEN D.P., TUDDENHAM E.G.;
 RT "Purification and characterization of factor VIII 1,689-Cys: a
 RT nonfunctional cofactor occurring in a patient with severe hemophilia
 RT A.";
 RL BLOOD 73:2117-2122(1989).
 RN [17]
 RP VARIANT CYS-391.
 RX MEDLINE; 90001543.
 RA SHIMA M., WARE J., YOSHIOKA A., FUKUI H., FULCHER C.A.;
 RT "An arginine to cysteine amino acid substitution at a critical
 RT thrombin cleavage site in a dysfunctional factor VIII molecule.";
 RL BLOOD 74:1612-1617(1989).
 RN [18]
 RP VARIANT LEU-189.
 RX MEDLINE; 90057680.
 RA CHAN V., CHAN T.K., TONG T.M., TODD D.;
 RT "A novel missense mutation in exon 4 of the factor VIII:C gene
 RT resulting in moderately severe hemophilia A.";
 RL BLOOD 74:2688-2691(1989).
 RN [19]
 RP VARIANT LEU-2326.
 RX MEDLINE; 89197216.
 RA INABA H., FUJIMAKI M., KAZAZIAN H.H. JR., ANTONARAKIS S.E.;
 RT "Mild hemophilia A resulting from Arg-to-Leu substitution in exon 26
 RT of the factor VIII gene";
 RL HUM. GENET. 81:335-338(1989).
 RN [20]
 RP VARIANT HIS-391.
 RX MEDLINE; 89264602.
 RA ARAI M., INABA H., HIGUCHI M., ANTONARAKIS S.E., KAZAZIAN H.H. JR.,
 RA FUJIMAKI M., HOYER L.W.;
 RT "Direct characterization of factor VIII in plasma: detection of a
 RT mutation altering a thrombin cleavage site
 RT (arginine-372-->histidine).";
 RL PROC. NATL. ACAD. SCI. U.S.A. 86:4277-4281(1989).
 RN [21]
 RP VARIANT CYS-1708.
 RX MEDLINE; 90105723.
 RA ARAI M., HIGUCHI M., ANTONARAKIS S.E., KAZAZIAN H.H. JR.,
 RA PHILLIPS J.A. III, JANCO R.L., HOYER L.W.;
 RT "Characterization of a thrombin cleavage site mutation (Arg 1689 to
 RT Cys) in the factor VIII gene of two unrelated patients with
 RT cross-reacting material-positive hemophilia A.";
 RL BLOOD 75:384-389(1990).
 RN [22]
 RP VARIANTS GLN-2228 AND LEU-2326.
 RX MEDLINE; 90123183.
 RA CASULA L., MURRU S., PECORARA M., RISTALDI M.S., RESTAGNO G.,
 RA MANCUSO G., MORFINI M., DE BIASI R., BAUDO F., CARONARA A.;
 RT "Recurrent mutations and three novel rearrangements in the factor
 RT VIII gene of hemophilia A patients of Italian descent";
 RL BLOOD 75:662-670(1990).
 RN [23]
 RP VARIANT CYS-391.
 RX MEDLINE; 90329422.
 RA PATTERSON J.K., MCVEY J.H., BOON M., AJANI A., TUDDENHAM E.G.;
 RT "CRM+ hemophilia A due to a missense mutation (372-->Cys) at the


```

RN  [2]
RP  SEQUENCE OF 170-387 FROM N.A.
RC  TISSUE-MAMMARY GLAND;
RX  MEDLINE: 91371351
RA  LAROCCA D., PETERSON J.A., URREA R., KUNIVOSHI J., BISTRAN A.M.,
RA  CERRANI R.L.;
RT  "A Mr 46,000 human milk fat globule protein that is highly expressed
RL  in human breast tumors contains factor VIII-like domains.";
RC  CANCER RES. 51:4994-4998(1991).
CC  -!- FUNCTION: MAY BE INVOLVED IN PHOSPHOLIPID BINDING.
CC  -!- SUBCELLULAR LOCATION: PERIPHERAL MEMBRANE PROTEIN.
CC  -!- TISSUE SPECIFICITY: MAMMARY EPITHELIAL CELL SURFACES.
CC  -!- OVEREXPRESSED IN SEVERAL CARCINOMAS.
CC  -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC  -!- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
CC  -!- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL: U58516; G1381162;
CC  EMBL: S56151; G235397;
CC  MIM: 602281;
CC  PROSITE: PS00022; EGF_1; 1.
CC  PROSITE: PS01186; EGF_2; 1.
CC  PROSITE: PS01285; FA58C_1; 2.
CC  PROSITE: PS01286; FA58C_2; 2.
CC  PFAM: PF00008; EGF; 1.
CC  PFAM: PF00754; F5_F8_type_C; 2.
CC  SIGNAL: GLYCOPROTEIN; MILK; REPEAT: EGF-LIKE DOMAIN.
CC  SIGNAL 1 22 POTENTIAL.
CC  CHAIN 23 387 MILK FAT GLOBULE-EGF FACTOR 8.
CC  DOMAIN 23 67 EGF-LIKE.
CC  DOMAIN 70 225 F5/8 TYPE C 1.
CC  DOMAIN 230 387 F5/8 TYPE C 2.
CC  SITE 46 48 CELL ATTACHMENT SITE (POTENTIAL).
CC  DISULFID 27 38 BY SIMILARITY.
CC  DISULFID 32 55 BY SIMILARITY.
CC  DISULFID 57 66 BY SIMILARITY.
CC  DISULFID 70 225 BY SIMILARITY.
CC  DISULFID 212 216 BY SIMILARITY.
CC  DISULFID 230 387 BY SIMILARITY.
CC  CARBOHYD 238 238 POTENTIAL.
CC  CARBOHYD 325 325 POTENTIAL.
CC  CARBOHYD 329 329 POTENTIAL.
CC  CARBOHYD 350 350 POTENTIAL.
CC  SEQUENCE 387 AA; 43123 MW; 9672347E CRC32;
SQ

```

```

Query Match 8.2%; Score 177.5; DB 1; Length 387;
Best Local Similarity 34.1%; Pred. No. 1.8e-07;
Matches 58; Conservative 27; Mismatches 58; Indels 27; Gaps 10;

QY 13 CRYALGNQRTIPDSDISASS-----SWSDSTAARHSRLSSDGDGAWCPAGSVF 62
DB 230 CANPLGKKNIPDKQITASSYKTKWGLHFSWNPFS---YARLDKQGNFAWV-AGSY- 284
QY 63 PKEEYQLVDLQRLHLVALVGTQG-RHAGLGKGFERSYRLRYSRDGRRWGKWD-RWG- 120
DB 284 -GNDWLQVLDLGSKEVTKGIITQGAENFGSV--QFVASYKVAISNDSANWTEYQDPRTGS 340
QY 120 QEVISGNEDEGVVLDLGPVAVLVRYFPA--DRVMSVCLRVLYGCG 167
DB 341 SKIPFGWONHSHKKNLFETPIILARYVRLPVAWNR----IALRLELLGC 387

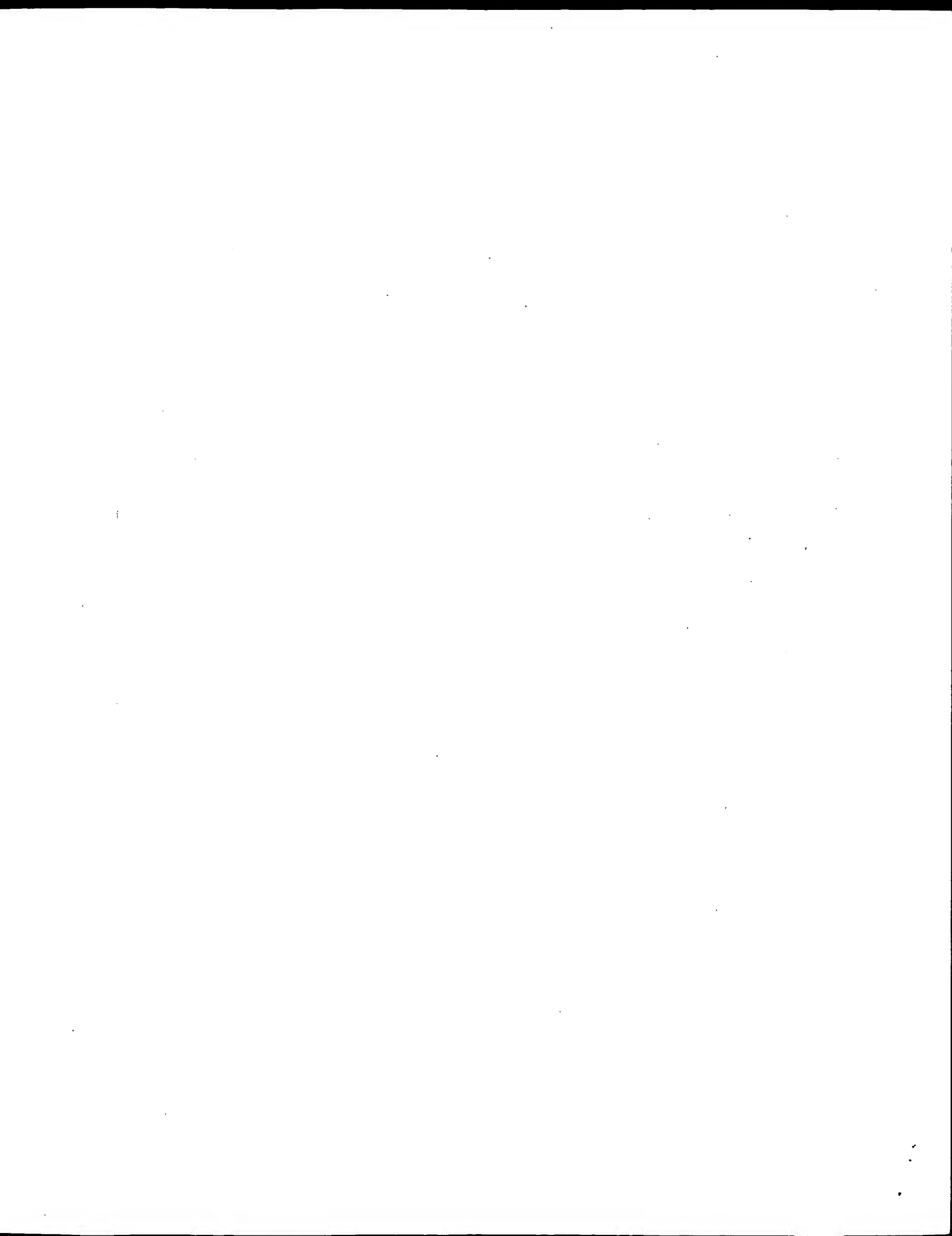
RESULT 13
FAS_MOUSE
ID FAS_MOUSE STANDARD; PRT: 2319 AA.

```

```

AC Q06194;
DT 01-JUN-1994 (REL. 29, CREATED)
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE COAGULATION FACTOR VIII PRECURSOR (PROCOAGULANT COMPONENT).
GN CF8 OR F8C.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C57BL/6 X CBA; TISSUE-LIVER;
RX MEDLINE: 93300511.
RA ELDERS B., LAKICH D., GITTSCHER J.;
RT "Sequence of the murine factor VIII cDNA.";
RL GENOMICS 16:374-379(1993).
CC -!- FUNCTION: FACTOR VIII, ALONG WITH CALCIUM AND PHOSPHOLIPID, ACTS
CC AS A COFACTOR FOR FACTOR IXA WHEN IT CONVERTS FACTOR X TO THE
CC ACTIVATED FORM, FACTOR XA.
CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -!- TISSUE SPECIFICITY: FOUND IN MOST TISSUES.
CC -!- SIMILARITY: CONTAINS 3 F5/8 TYPE A DOMAINS; EACH IS COMPOSED OF
CC 2 PLASTOCYANIN-LIKE REPEATS.
CC -!- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
CC -!- SIMILARITY: STRONG, TO COAGULATION FACTOR V.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC  -----
CC EMBL: L05573; G192457;
CC PIR: A47004; A47004.
CC MGD; MGI:88383; CF8.
CC PROSITE: PS00079; MULTICOPPER_OXIDASEL; 3.
CC PROSITE: PS01285; FA58C_1; 2.
CC PROSITE: PS01286; FA58C_2; 2.
CC PFAM: PF00394; Cu-oxidase; 3.
CC PFAM: PF00754; F5_F8_type_C; 2.
CC HSP: P00451; ICFG.
CC BLOOD COAGULATION; REPEAT: PLASMA; ACUTE PHASE; CALCIUM;
KW SIGNAL; GLYCOPROTEIN; SULFATATION.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 2319 COAGULATION FACTOR VIII.
FT DOMAIN 20 349 F5/8 TYPE A 1.
FT DOMAIN 20 199 PLASTOCYANIN-LIKE 1.
FT DOMAIN 207 349 PLASTOCYANIN-LIKE 2.
FT DOMAIN 399 730 F5/8 TYPE A 2.
FT DOMAIN 399 573 PLASTOCYANIN-LIKE 3.
FT DOMAIN 583 730 PLASTOCYANIN-LIKE 4.
FT DOMAIN 760 1640 B.
FT DOMAIN 1683 2008 F5/8 TYPE A 3.
FT DOMAIN 1683 1845 PLASTOCYANIN-LIKE 5.
FT DOMAIN 1855 2008 PLASTOCYANIN-LIKE 6.
FT DOMAIN 2008 2156 F5/8 TYPE C 1.
FT DOMAIN 2161 2313 F5/8 TYPE C 2.
FT SITE 391 392 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT SITE 759 760 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT SITE 1678 1679 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT SITE 1324 1325 CLEAVAGE (ACTIVATION) (BY SIMILARITY).
FT SITE 1640 1641 CLEAVAGE (ACTIVATION) (BY SIMILARITY).
FT MOD_RES 367 367 SULFATATION (BY SIMILARITY).
FT MOD_RES 737 737 SULFATATION (BY SIMILARITY).
FT MOD_RES 738 738 SULFATATION (BY SIMILARITY).
FT MOD_RES 742 742 SULFATATION (BY SIMILARITY).
FT MOD_RES 1669 1669 SULFATATION (REQUIRED FOR VWF BINDING)
FT MOD_RES 1687 1687 (BY SIMILARITY).
FT DISULFID 173 199 SULFATATION (BY SIMILARITY).
FT DISULFID 547 573 PROBABLE.

```

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 4, 1999, 01:52:29 ; Search time 19.79 Seconds
(without alignments)
807.788 Million cell updates/sec

Title: US-08-170-558-8
Perfect score: 2175
Sequence: 1 DADMGHFDPAKCRYALGMQ.....LELEPRGQOPVAKPEGSPTA 399

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database : PIR_60.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2167	99.6	876	2 A49508	protein-tyrosine k
2	2164	99.5	913	2 A48280	receptor tyrosine
3	2021	92.9	910	2 A53137	tyrosine kinase re
4	1002.5	46.1	855	2 S42621	protein-tyrosine k
5	952.5	43.8	819	2 I48859	tyro 10 receptor k
6	213	9.8	927	2 JQ0948	A5 antigen precurs
7	211.5	9.7	845	3 JC5256	adipocyte transcri
8	192	8.8	2351	1 E2HU	coagulation factor
9	192	8.8	216	2 A44258	factor VIII-associ
10	179.5	8.3	2224	1 KF0U5	coagulation factor
11	177.5	8.2	218	2 A47285	milk fat globule p
12	175	8.0	2319	2 A47004	coagulation factor
13	174.5	8.0	2211	1 KF0U5	coagulation factor
14	174.5	8.0	463	2 A36479	milk fat globule m
15	172.5	7.9	427	2 JC4915	ags protein precur
16	167.5	7.7	401	2 S65138	glycoprotein antig
17	167.5	7.7	427	2 S74211	PAS-6/7 protein pr
18	164.5	7.6	719	2 S51739	transcription repr
19	145.5	6.7	3133	2 S52093	hemocytin - silkwo
20	106.5	4.9	448	2 A57435	3,4-dihydroxypheny
21	97	4.5	253	1 DL001A	discoicoidin I chain
22	97	4.5	1269	2 S35366	furl protein - fru
23	97	4.5	899	2 S17546	probable serine pr
24	96	4.4	253	1 DL001C	discoicoidin I chain
25	94.5	4.3	395	2 S76793	hypothetical prote
26	92	4.2	635	2 S74718	sulfite reductase
27	89	4.1	3164	1 WMBEH6	UL36 protein - hum
28	89	4.1	477	2 S2162	sucrose hydrolase
29	88	4.0	1384	3 T02748	hypothetical prote
30	87.5	4.0	139	1 C45051	lamprin 2 precurs
31	87.5	4.0	1168	1 MWAXIC	myosin heavy chain
32	87.5	4.0	420	2 A70796	hypothetical prote
33	87	4.0	439	2 I60916	HNF-3/forhead hom
34	87	4.0	404	2 S66533	flavoprotein A, 45
35	87	4.0	409	2 E69046	flavoprotein AI -
36	86.5	4.0	149	1 DL001D	discoicoidin I chain
37	86.5	4.0	377	2 A45851	MHC class I histoc
38	86.5	4.0	149	2 B03382	discoicoidin I chain
39	86.5	4.0	379	2 S09791	hypothetical prote

ribosomal protein
probable regulator
DNA-directed RNA p
hypothetical prote
Bicaudal-C - fruit
hypothetical prote

40 86 4.0 253 2 S53847
41 86 4.0 1085 2 E70834
42 85.5 3.9 1460 1 RNBY3L
43 85.5 3.9 323 2 T02617
44 85.5 3.9 938 2 S55051
45 85.5 3.9 458 2 SI2444

ALIGNMENTS

RESULT 1
A49508
protein-tyrosine kinase (EC 2.7.1.112) trke precursor - human
C:Species: Homo sapiens (man)
C:Date: 10-Nov-1995 #sequence.revision 10-Nov-1995 #text_change 16-Dec-1998
C:Accession: A49508; I38358; S37402
R:Di Marco, E.; Cutuli, N.; Guerra, L.; Cancedda, R.; De Luca, M.
J. Biol. Chem. 288, 24290-24295, 1993
A:Title: Molecular cloning of trke, a novel trk-related putative tyrosine kinase rece
A:Reference number: A49508; MUID:94043265
A:Accession: A49508
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-876 <DIA>
A:Cross-references: EMBL:X74979; NID:g400462; PID:g400463
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; discoidin I amino-
C:Keywords: Anp; phosphotransferase
F:30-185/Domain: discoidin I amino-terminal homology <DNI>
F:571-875/Domain: protein kinase homology <KIN>
F:579-587/Region: protein kinase ATP-binding motif

Query Match 99.6% Score 2167; DB 2; Length 876;
Best Local Similarity 99.7%; Pred. No. 2.2e-177;
Matches 398; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 DADMGHFDPAKCRYALGMQDRTIPDSISASSWSWSDTAARHSRLSSDGDGAWCPAGS 60
Db 19 DADMGHFDPAKCRYALGMQDRTIPDSISASSWSWSDTAARHSRLSSDGDGAWCPAGS 78
Oy 61 VFPKEEYLVQDLQRLHLVALVTQGRHAGLGKEFSSRYRLYSRDRGRMWGKDRWGQ 120
Db 79 VFPKEEYLVQDLQRLHLVALVTQGRHAGLGKEFSSRYRLYSRDRGRMWGKDRWGQ 138
Oy 121 EVISGNEDEPGVVLKDLGPPMVARLVRFYPRADRVMSVCLRVLYGCLWRDGLLSYAPV 180
Db 139 EVISGNEDEPGVVLKDLGPPMVARLVRFYPRADRVMSVCLRVLYGCLWRDGLLSYAPV 198
Oy 181 GQTYLSEAVYVNDSTYDGHVTGGLQYGGGLQGLADGVVGLDDFRKSOELRWMPGYDYVGW 240
Db 199 GQTYLSEAVYVNDSTYDGHVTGGLQYGGGLQGLADGVVGLDDFRKSOELRWMPGYDYVGW 258
Oy 241 SNHFSFGVYEMEFDFRLRAFAQMVHCNMHTLGLARLPGGVECFRRFRPGPAMAWEGEPM 300
Db 259 SNHFSFGVYEMEFDFRLRAFAQMVHCNMHTLGLARLPGGVECFRRFRPGPAMAWEGEPM 318
Oy 301 RHNLGNGLDPRARVSVPLGGVRARELQCRFLFAGPWLFLSEISFISDVYNNSSPALGG 360
Db 319 RHNLGNGLDPRARVSVPLGGVRARELQCRFLFAGPWLFLSEISFISDVYNNSSPALGG 378
Oy 361 TFPAPMWPMPGPPPTNFSSLELEPRGQOPVAKPEGSPTA 399
Db 379 TFPAPMWPMPGPPPTNFSSLELEPRGQOPVAKPEGSPTA 417

RESULT 2
A48280
receptor tyrosine kinase - human
C:Species: Homo sapiens (man)
C:Date: 31-May-1996 #sequence.revision 31-May-1996 #text_change 16-Dec-1998
C:Accession: A48280
R:Johnson, J.D.; Edman, J.C.; Rutter, W.J.

Proc. Natl. Acad. Sci. U.S.A. 90, 5677-5681, 1993

A:Title: A receptor tyrosine kinase found in breast carcinoma cells has an extracellular

A:Reference number: A48280; MUID:93296201

A:Accession: A48280

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-913 <RES>

A:Cross-references: GB:L11315; NID:g403386; PID:g403387

C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases

F:30-185/Domain: discoidin I amino-terminal homology <DN1>

F:608-912/Domain: protein kinase homology <KIN>

F:616-624/Region: protein kinase ATP-binding motif

Query Match 99.5%; Score 2164; DB 2; Length 913;

Best Local Similarity 99.5%; Pred. No. 4.2e-177;

Matches 397; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DADMGHFDPAKRYALGMQDRTIPDSISASSWSSTAARHSRLSSDGDGAWCPAGS 60

DB 19 DADMGHFDPAKRYALGMQDRTIPDSISASSWSSTAARHSRLSSDGDGAWCPAGS 78

QY 61 VFPKEEYLVQDLQRLHLVALVGTQGRHAGGLGKEFSYRLYSRDGRRWGKDRWGQ 120

DB 79 VFPKEEYLVQDLQRLHLVALVGTQGRHAGGLGKEFSYRLYSRDGRRWGKDRWGQ 138

QY 121 EVISNEDPEGVVLKDLGPPMVARLVREYPRADRVMSVCLRVLYGCLWRDGLLSYAPV 180

DB 139 EVISNEDPEGVVLKDLGPPMVARLVREYPRADRVMSVCLRVLYGCLWRDGLLSYAPV 198

QY 181 GOTMYLSEAVYLNDSYDCHTVGGYGLGQLADGVVGLDFFRKSOELRWMPGYDYVGW 240

DB 199 GOTMYLSEAVYLNDSYDCHTVGGYGLGQLADGVVGLDFFRKSOELRWMPGYDYVGW 258

QY 241 SNHSSFGYVEMEFEDRLRAFOAMOVHCNNMHTLGARLPGVCECRFRGPPAMAWECEPM 300

DB 259 SNHSSFGYVEMEFEDRLRAFOAMOVHCNNMHTLGARLPGVCECRFRGPPAMAWECEPM 318

QY 301 RHNLGGLGDPARAVSVPLGGRVARFLQCRFLFAGPWLFLSEISFISDVYNNSSPALGG 360

DB 319 RHNLGGLGDPARAVSVPLGGRVARFLQCRFLFAGPWLFLSEISFISDVYNNSSPALGG 378

QY 361 TFPAPMPWPPGPPPTNFSSLELEPRGQOPVAKPEGSPPTA 399

DB 379 TFPAPMPWPPGPPPTNFSSLELEPRGQOPVAKPEGSPPTA 417

RESULT 3

A53137

tyrosine kinase receptor - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 16-Dec-1998

C:Accession: A53137

R:Sanchez, M.P.; Tapley, P.; Saini, S.S.; He, B.; Pulido, D.; Barbacid, M.

Proc. Natl. Acad. Sci. U.S.A. 91, 1819-1823, 1994

A:Title: Multiple tyrosine protein kinases in rat hippocampal neurons: isolation of Ptk-

A:Reference number: A53137; MUID:94173920

A:Accession: A53137

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-910 <RES>

A:Cross-references: GB:L26525; NID:g432480; PID:g432481

C:Genetics:

A:Gene: Ptk-3

C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; discoidin I amino-ter

F:31-186/Domain: discoidin I amino-terminal homology <DN1>

F:605-909/Domain: protein kinase homology <KIN>

F:613-621/Region: protein kinase ATP-binding motif

Query Match

Best Local Similarity 92.9%; Score 2021; DB 2; Length 910;

Matches 371; Conservative 9; Mismatches 15; Indels 4; Gaps 1;

QY 1 DADMGHFDPAKRYALGMQDRTIPDSISASSWSSTAARHSRLSSDGDGAWCPAGS 60

DB 20 DADMGHFDPAKRYALGMQDRTIPDSISASSWSSTAARHSRLSSDGDGAWCPAGP 79

QY 61 VFPKEEYLVQDLQRLHLVALVGTQGRHAGGLGKEFSYRLYSRDGRRWGKDRWGQ 120

DB 80 VFPKEEYLVQDLQRLHLVALVGTQGRHAGGLGKEFSYRLYSRDGRRWGKDRWGQ 139

QY 121 EVISNEDPEGVVLKDLGPPMVARLVREYPRADRVMSVCLRVLYGCLWRDGLLSYAPV 180

DB 140 EVISNEDPEGVVLKDLGPPMVARLVREYPRADRVMSVCLRVLYGCLWRDGLLSYAPV 199

QY 181 GOTMYLSEAVYLNDSYDCHTVGGYGLGQLADGVVGLDFFRKSOELRWMPGYDYVGW 240

DB 200 GOTMYLSEAVYLNDSYDCHTVGGYGLGQLADGVVGLDFFRKSOELRWMPGYDYVGW 259

QY 241 SNHSSFGYVEMEFEDRLRAFOAMOVHCNNMHTLGARLPGVCECRFRGPPAMAWECEPM 300

DB 260 SNHSSFGYVEMEFEDRLRAFOAMOVHCNNMHTLGARLPGVCECRFRGPPAMAWECEPM 319

QY 301 RHNLGGLGDPARAVSVPLGGRVARFLQCRFLFAGPWLFLSEISFISDVYNNSSPALGG 360

DB 320 RHNLGGLGDPARAVSVPLGGRVARFLQCRFLFAGPWLFLSEISFISDVYNNSSPALGG 375

QY 361 TFPAPMPWPPGPPPTNFSSLELEPRGQOPVAKPEGSPPTA 399

DB 376 TFPAPMPWPPGPPPTNFSSLELEPRGQOPVAKPEGSPPTA 414

RESULT 4

S42621

protein-tyrosine kinase (EC 2.7.1.112) - human

C:Species: Homo sapiens (man)

C:Date: 07-Sep-1994 #sequence_revision 10-Nov-1995 #text_change 17-Mar-1999

C:Accession: S42621

Oncogene 8, 3433-3440, 1993

A:Title: Structure, expression and chromosomal mapping of TKT from man and mouse: a n

A:Reference number: S42621; MUID:94067796

A:Accession: S42621

A:Molecule type: mRNA

A:Residues: 1-855 <KAR>

A:Cross-references: EMBL:X74764; NID:g433337; PID:g433338

C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases

C:Keywords: phosphotransferase; tyrosine-specific protein kinase

F:29-185/Domain: discoidin I amino-terminal homology <DN1>

F:561-855/Domain: protein kinase homology <KIN>

F:569-577/Region: protein kinase ATP-binding motif

Query Match 46.1%; Score 1002.5; DB 2; Length 855;

Best Local Similarity 53.6%; Pred. No. 8.1e-78;

Matches 196; Conservative 54; Mismatches 109; Indels 7; Gaps 6;

QY 5 KGHFDPAKRYALGMQDRTIPDSISASSWSSTAARHSRLSSDGDGAWCPAGSVPK 64

DB 22 KAQVNPACRYPLGSGGQIPDEDITASSWSESTAAYGRDSEEGDGAWCPEIPVEPD 81

QY 65 E-EYLVQDLQRLHLVALVGTQGRHAGGLGKEFSYRLYSRDGRRWGKDRWGQEV 123

DB 82 DLKEFLQIDLHTLHTITLVGTQGRHAGGHGIEFAPMYKINYSRDGTRWISWNRHGKQVL 141

QY 124 SGNEDPEGVVLKDLGPPMVARLVREYPRADRVMSVCLRVLYGCLWRDGLLSYAPV 183

DB 142 DGNNSNPIDFLKDLLEPIVARVRIPIVTDHSMVMCMRVLYGCVLWDLGLSYNAPAQ 201

QY 184 MYL--SEAVYLNDSYDCHTVGGYGLGQLADGVVGLDFFRKSOELRWMPGYDYVGW 241

DB 202 FVLPGSIIYLNDSYVDG-AVGYSMTGEGLGQLTDGVSGLDDFTQTHEYHVMPGYDYVGW 260

QY 242 NISFSSGYVEMEFEDRLRAFOAMOVHCNNMHTLGARLPGVCECRFRGPPAMAWECEPM 301

Db 261 NESATNGYIEIMFEPRDRINFTTMKVHCNNMFAKGVKIFKVCQYF-RSEASEWEPNPAIS 319
QY 302 HNLGGNLGDPARAVSVPLGGVRAREFLOCRFLFAGPWLLEFSEITSD-VVNNSSPALGG 360
Db 320 FPLVLDVNPASREFTVPLHHRMASAIKQYHFAETWMMFSEITFQSDAAMYNSEAL-P 378
QY 361 TFPAP 366
Db 379 TSPMAP 384

RESULT 5
I48859
cyto 10 receptor kinase - mouse
C:Species: Mus musculus (house mouse)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 19-Dec-1997
C:Accession: I48859; 842796
R:Lai, C.; Lemke, G.
Oncogene 9, 877-883, 1994
A:Title: Structure and expression of the Tyro 10 receptor tyrosine kinase.
A:Reference number: I48859; MUID:94151011
A:Accession: I48859
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-819 <RES>
A:Cross-references: EMBL:X76505; NID:g435161; PID:g435162
C:Genetics:
A:Gene: tyro 10
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
C:Keywords: ATP
F:526-819/Domain: protein kinase homology <KIN>
F:534-542/Region: protein kinase ATP-binding motif

Query Match 43.8%; Score 952.5; DB 2; Length 819;
Best Local Similarity 49.7%; Pred. No. 1.4e-73;
Matches 189; Conservative 57; Mismatches 107; Indels 27; Gaps 7;

QY 24 IPDSISASSWSSTAAARHSRLSSDGDGAWCPAGSVFPKE-EYQLVDLQRLHLVALV 82
Db 6 IPDEDITASSQWSESTAAKYGRLDSEEGDGAWCPEIPQPDLDLKEFLQIDRLTHFLIV 65

QY 83 GTQGRAGGLGFEFSYRLYSRGRMMGWKDRWGQVVISGNEDEPGVVLKDLGPPMV 142
Db 66 GTQGRAGGAGGIEFAPMYKINISRDGSRWISWRNHRGKQVLDGNSNPYDFLKLPEPIV 125

QY 143 ARLVREYPRADRVMSVCLAVELYGCLWRDGLLSYAPYQOTWYL--SEAVYLNDSTYDGH 200
Db 126 ARFVRLIPTDHSNMVCMRVELYGCVLWDLGVSYNAPAGQGVLPGGSIYIYNDSSVYDG- 185

QY 201 TVGGLQYGLGQADGVVGLDFFRKSQELRVWPGYDYGVWSNHSFSSGVVMEFEFFDLRL 260
Db 185 AVGYSMTGLGQLTDCVSGLDLDFQTQTHEYHVWPGYDYGVWRNESATNGFIEIMFEEDRIR 244

QY 261 AFQAMQYHNNHTLIGARLPGGVECFRGRPMAGEPEMRINLGNLGDPRARAVSVPL 320
Db 245 NETTMKVHCNNMFAKGVKIFKVCQYF-RSEASEWEPNPAISREFTVPL 303

QY 321 GGVARFLOCRFLFAGPWLLEFSEITSD--VVNNSSPALGTFPPAPWPGPPPTNFS 378
Db 304 HHRMASAIKQYHFAETWMMFSEITFQSDAAMYNSS--GALPTSP----- 347

QY 379 SLELEPRGOQPVAKPEGSPT 398
Db 347 ---MAPTTYDPMKLKVDSDNT 363

RESULT 6
JQ0948
A5 antigen precursor - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C>Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 20-Mar-1998
C:Accession: JH0466; JQ0948

R:Takagi, S.; Hirata, T.; Agata, K.; Mochii, M.; Eguchi, G.; Fujisawa, H.
Neuron 7, 295-307, 1991
A:Title: The A5 antigen, a candidate for the neuronal recognition molecule, has homol
A:Reference number: JH0466; MUID:91337458
A:Accession: JH0466
A:Molecule type: mRNA
A:Residues: 1-927 <TAK>
A:Cross-references: GB:D10467; GB:D01077; NID:g222962; PID:d1001730; PID:g222963
A:Experimental source: tadpole, brain
A:Note: This protein has motifs homologous to complement components C1r and C1s and t
C:Comment: This protein is a neuronal cell surface molecule involved in the neuronal
C:Superfamily: Xenopus A5 antigen; C1r/C1s repeat homology; discoidin I amino-termina
C:Keywords: duplication; glycoprotein; transmembrane protein
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-927/Product: A5 antigen #status predicted <A5A>
F:27-138/Domain: C1r/C1s repeat homology <C1R1>
F:147-262/Domain: C1r/C1s repeat homology <C1R2>
F:274-424/Domain: discoidin I amino-terminal homology <DN1>
F:430-584/Domain: discoidin I amino-terminal homology <DN2>
F:646-812/Domain: MAM homology <MAM>
F:861-883/Domain: transmembrane #status predicted <TM>
F:150,261,300,523,844/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.8%; Score 213; DB 2; Length 927;
Best Local Similarity 33.3%; Pred. No. 2.9e-10;
Matches 60; Conservative 25; Mismatches 69; Indels 26; Gaps 6;

QY 12 KCRYALGMQDRTTPDSDISASSWSSTAAARHSRLSSDGDGAWCPAGSVFPKEEYQLV 71
Db 274 QCKEALGMSEGEIHFQISVSSQYSMNWSAERSRLNYV--ENGWTPGEDTV---KEWIOV 328

QY 72 DLQRLHLVALVGTQGRHAGLGKE----FSRSYRLYSRDRMMGWKDRWGQEVISGN 126
Db 329 DLENLRFVSGIGTQ----GAISRETKKKYFKSIKIDISNGEDWITLKDGNKHLVFTGN 384

QY 127 EDPEGVVLKDLGPPMVARLVRYPRADRVMSVCLRVELYGCLWRD-----GLIS 175
Db 385 TDATDVVYRPFSPVITRFVRLRP-VTWENGISLRFELYGCKITDYPCSRMLGMVSGSLIS 443

RESULT 7
JC5256
adipocyte transcription factor, AEBP1 - Human
C:Species: Homo sapiens (man)
C>Date: 04-Dec-1998 #sequence_revision 04-Dec-1998 #text_change 04-Dec-1998
C:Accession: JC5256
R:Ohno, I.; Hashimoto, J.; Shimizu, K.; Takaoaka, K.; Ochi, T.; Matsubara, K.; Okubo,
Biochem. Biophys. Res. Commun. 228, 411-414, 1996
A:Title: A cDNA cloning of human AEBP1 from primary cultured osteoblasts and its expr
A:Reference number: JC5256; MUID:97079196
A:Accession: JC5256
A:Status: preliminary
A:Residues: 1-845 <OHN>
A:Cross-references: DDBJ:D86479

Query Match 9.7%; Score 211.5; DB 3; Length 845;
Best Local Similarity 24.3%; Pred. No. 3.5e-10;
Matches 80; Conservative 54; Mismatches 144; Indels 51; Gaps 12;

QY 9 DPAKCRYALGMQDRTTPDSDISASSWSSTAAARHSRLSSDGD-----DGAWCPAGSV 61
Db 68 EKVYKCP-PIGMESHRIEDNQIRASSMLRHCLGAQRGLNMOTGATEDDYDGAWCAEDDA 126

QY 62 FPKEEYQLVDLQRLHLVALVGTQGRHAGLGKEFSRSYRLYSRDRMMGWKDRWGQ 121
Db 127 ---RTQWIEVDTRTRTFTGVTQGRDS-SIHDDFTVTFVGFSGNSQSTWVMYNGYEEM 182

QY 122 VISGNEDEPGVVLKDLGPPMVARLVRYPRADRVMSVCLRVELYGCLWRDGLLSYAPV 181
Db 183 TFGHNVKDQDPVLSELPVAVFIRIYPLTNV-GSLCMRLEVIGC-----SVAPVY 233

F:1740-1741/Cleavage site: Arg-Ala (coagulation factor Xa) #status experimental
F:2193-2345/Disulfide bonds: #status predicted

Query Match 8.8% Score 192; DB 1; Length 2351;
Best Local Similarity 31.88; Pred. No. 5.9e-08;
Matches 54; Conservative 32; Mismatches 68; Indels 16; Gaps 5;

QY 9 DPAKRYALGMDRTIPDSISASS-----SWSDDTAARHSRLSSDGDGAWCPAGSV 61
DB 2189 DLNCSMPLGMEKSAISDAITASSYTNMFATWSPSKARLHLOGRSN-----AWRFQVN- 2244
QY 62 FPKREEYLVLDLRLHLVALVGTQGRHAGGLGKEFFSYRLYSRDRGRWMGKDRWGQE 121
DB 2244 --NPKEWLVDFQTKMVTGTTQGVKS--LLTSMYVKEFLISSQDGHQWTLFFQNGKVK 2300
QY 122 VISGNEDEPGVWLKDLGPPVARLVREYPRADRVMSVCLRVLYGCLWRD 171
DB 2301 VFQGNQDSFTPVNSLDPLLTRILRHPOS--WVHQIALRMEVLGCEAQD 2349

RESULT 9
A44258
factor VIII-associated gene B hypothetical protein - human
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 31-Oct-1997
C:Accession: A44258
R:Levinson, B.; Kenwright, S.; Gamel, P.; Fisher, K.; Gitschler, J.
Genomics 14, 585-589, 1992
A:Title: Evidence for a third transcript from the human factor VIII gene.
A:Reference number: A44258; MUID:93052386
A:Accession: A44258
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-216 <LEV>
A:Cross-references: GB:M90707; NID:g182316; PID:g182317
A:Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxidase
F:1-53/Domain: discoidin I amino-terminal homology #status atypical <DN1>
F:57-210/Domain: discoidin I amino-terminal homology <DN2>

Query Match 8.8% Score 192; DB 2; Length 216;
Best Local Similarity 31.88; Pred. No. 2.8e-09;
Matches 54; Conservative 32; Mismatches 68; Indels 16; Gaps 5;

QY 9 DPAKRYALGMDRTIPDSISASS-----SWSDDTAARHSRLSSDGDGAWCPAGSV 61
DB 54 DLNCSMPLGMEKSAISDAITASSYTNMFATWSPSKARLHLOGRSN-----AWRFQVN- 109
QY 62 FPKREEYLVLDLRLHLVALVGTQGRHAGGLGKEFFSYRLYSRDRGRWMGKDRWGQE 121
DB 109 --NPKEWLVDFQTKMVTGTTQGVKS--LLTSMYVKEFLISSQDGHQWTLFFQNGKVK 165
QY 122 VISGNEDEPGVWLKDLGPPVARLVREYPRADRVMSVCLRVLYGCLWRD 171
DB 166 VFQGNQDSFTPVNSLDPLLTRILRHPOS--WVHQIALRMEVLGCEAQD 214

RESULT 10
KFHUS
coagulation factor V precursor - human
N:Alternate names: coagulation labile factor; proaccelerin
C:Species: Homo sapiens (man)
C:Date: 19-May-1989 #sequence_revision 02-Jun-1995 #text_change 31-Oct-1997
C:Accession: A56172; A42344; A28028; A27498; A25897
R:Cripe, L.D.; Moore, K.D.; Kane, W.H.
Biochemistry 31, 3777-3785, 1992
A:Title: Structure of the gene for human coagulation factor V.
A:Reference number: A42344; MUID:92232668
A:Accession: A56172
A:Molecule type: DNA
A:Residues: 1-2224 <CRT>
A:Cross-references: GB:J05368

A:Accession: A42344
A:Molecule type: DNA
A:Residues: 48-58,79-89;120-130;191-201;239-249;313-323;368-378;428-437;461-471;533-5
2070;2111-2120;2172-2181 <CRZ>
R:Jenny, R.J.; Pittman, D.D.; Toole, J.J.; Kriz, R.W.; Aldape, R.A.; Hewick, R.M.; Ka
Proc. Natl. Acad. Sci. U.S.A. 84, 4846-4850, 1987
A:Title: Complete cDNA and derived amino acid sequence of human factor V.
A:Reference number: A28028; MUID:87260886
A:Accession: A28028
A:Molecule type: mRNA
A:Residues: 1-857, 'R', 859-864, 'R', 866-924, 'E', 926-1763, 'I', 1765-2212, 'T', 2214-2224 <J
A:Cross-references: GB:M16967
A:Note: parts of this sequence, including the amino end of the mature protein, were d
R:Kane, W.H.; Ichinose, A.; Hagen, F.S.; Davie, E.W.
Biochemistry 26, 6508-6514, 1987
A:Title: Cloning of cDNAs coding for the heavy chain region and connecting region of
A:Reference number: A27498; MUID:88107560
A:Accession: A27498
A:Molecule type: mRNA
A:Residues: 1-1284, 'I', 1286-1600 <KAN>
A:Cross-references: GB:M17785
A:Note: parts of this sequence were determined by protein sequencing
R:Kane, W.H.; Davie, E.W.
Proc. Natl. Acad. Sci. U.S.A. 83, 6800-6804, 1986
A:Title: Cloning of a cDNA coding for human factor V, a blood coagulation factor homo
A:Reference number: A25897; MUID:86313665
A:Accession: A25897
A:Molecule type: mRNA
A:Residues: 1188-1215, 1315-2224 <KA2>
A:Cross-references: GB:M14335
A:Note: parts of this sequence were determined by protein sequencing
R:Keller, F.G.; Ortel, T.L.; Quinn-Alten, M.A.; Kane, W.H.
Biochemistry 34, 4118-4124, 1995
A:Title: Thrombin-catalyzed activation of recombinant human factor V.
A:Reference number: A56139; MUID:95210278
A:Contents: annotation; thrombin cleavage sites
C:Comment: Factor V is activated by thrombin and partially by coagulation factor Xa.
C:Genetics:
A:Gene: GDB:F5
A:Cross-references: GDB:l19896; OMIM:227400
A:Map position: 1q23-1q23
A:Introns: 53/2; 84/1; 125/1; 196/1; 244/1; 318/1; 373/2; 432/3; 466/1; 537/3; 588/1;
C:Function:
A:Description: acts as a cofactor, with calcium and phospholipid, for the factor Xa p
A:Pathway: blood coagulation
C:Superfamily: coagulation factor V; discoidin I amino-terminal homology; ferroxidase
C:Keywords: blood coagulation; duplication; glycoprotein; phospholipid binding; plasm
F:1-28/Domain: signal sequence #status predicted <SIG>
F:29-2224/Product: coagulation factor V #status predicted <MAT>
F:29-737/Product: coagulation factor Va heavy chain #status experimental <VAH>
F:29-345/Domain: A1 <DA1>
F:33-329/Domain: ferroxidase repeat homology <FO1>
F:346-691/Domain: A2 <DA2>
F:351-684/Domain: ferroxidase repeat homology <FO2>
F:692-1573/Domain: B <DOB>
F:1183-1461/Region: 9-residue repeats (O-X-T/N-L-S-P-D-L-S)
F:1574-2224/Product: coagulation factor Va light chain #status experimental <VAL>
F:1574-1905/Domain: A3 <DA3>
F:1581-1905/Domain: ferroxidase repeat homology <FO3>
F:1667-1765/Region: phospholipid binding #status predicted
F:1906-2064/Domain: C1 <DC1>
F:1906-2061/Domain: discoidin I amino-terminal homology <DN1>
F:2065-2224/Domain: C2 <DC2>
F:2065-2221/Domain: discoidin I amino-terminal homology <DN2>
F:51-55,239,297,460,468,554,741,752,760,776,782,821,938,977,1074,1083,1103,1106,1479,
F:167-193,248-329,500-526,603-684,1725-1751,1907-2061,2066-2221/Disulfide bonds: #sta
F:334-335/Cleavage site: Arg-Asn (protein C) #status predicted
F:363,693,1546/Binding site: sulfate (Tyr) (covalent) #status predicted
F:376-377/Cleavage site: Arg-Ser (coagulation factor Xa) #status predicted
F:382,1338/Binding site: carboxylate (Asn) (covalent) #status absent
F:534-535/Cleavage site: Arg-Gly (protein C) #status predicted
F:737-738/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experiment
F:1046-1047/Cleavage site: Arg-Thr (coagulation factor Xa, thrombin) #status experime

```
F:1-19/Domain: signal sequence #status predicted <sig>
```

F1693-2046/DOMAIN: DISCOLDIN 1 AMINO-TERMINAL HOMOLOGY **NCNT2**

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run On: November 4, 1999, 01:52:00 ; Search time 18.36 Seconds
(without alignments)
248.367 Million cell updates/sec

Title: US-08-170-558-8
Perfect score: 2175
Sequence: 1 DADMKGHFDPKCRYALGMQ.....LELEPRGQPPVAKPEGSPTA 399

Scoring table: BLOSUM62

Searched: 119832 seqs, 11428610 residues

Database : Issued Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/PCTUS9_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2175	100.0	913	1	US-08-445-640-4
2	2175	100.0	399	1	US-08-445-640-8
3	2167	99.6	919	1	US-08-336-343A-2
4	1002.5	46.1	855	1	US-08-336-343A-4
5	995.5	45.8	854	2	US-08-456-647B-20
6	995.5	45.8	854	2	US-08-237-401A-20
7	210	9.7	2133	2	US-08-670-707A-37
8	210	9.7	1443	2	US-08-670-707A-39
9	201.5	9.3	1128	1	US-08-111-939-2
10	193	8.9	218	1	US-07-607-538C-5
11	192	8.8	2332	1	US-07-864-004B-4
12	192	8.8	2351	1	US-08-121-202-2
13	192	8.8	2332	1	US-08-251-937A-4
14	192	8.8	2332	1	US-08-212-133A-2
15	192	8.8	2351	1	US-08-366-831A-2
16	192	8.8	2332	1	US-08-276-594A-2
17	192	8.8	1471	1	US-08-683-839B-3
18	192	8.8	2332	1	US-08-474-503-2
19	192	8.8	2332	1	US-08-670-707A-2
20	192	8.8	1661	2	US-08-882-083-2
21	192	8.8	1661	2	US-08-558-107-2
22	192	8.8	2332	3	PCT-US93-03275-4
23	192	8.8	2332	3	PCT-US94-13200-2
24	179.5	8.3	218	1	US-07-607-538C-4
25	177.5	8.2	218	1	US-07-607-538C-2
26	177.5	8.2	217	1	US-07-607-538C-3
27	175	8.0	2319	1	US-08-212-133A-8
28	175	8.0	2319	1	US-08-474-503-6
29	175	8.0	2319	2	US-08-670-707A-6
30	175	8.0	2319	3	PCT-US94-13200-6
31	171	7.9	480	2	US-08-480-229C-10
32	171	7.9	480	2	US-08-659-235C-10
33	166.5	7.7	513	2	US-08-480-229C-14
34	166.5	7.7	513	2	US-08-659-235C-14
35	163.5	7.5	320	2	US-08-480-229C-20
36	163.5	7.5	320	2	US-08-659-235C-20
37	158.5	7.3	321	2	US-08-480-229C-21
38	158.5	7.3	321	2	US-08-659-235C-21
39	147.5	6.8	109	1	US-08-111-939-19

Sequence 23, Appl
Sequence 21, Appl
Sequence 25, Appl
Sequence 20, Appl
Sequence 24, Appl
Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-08-445-640-4
; Sequence 4, Application US/08445640
; Patent No. 5709858
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J.
; APPLICANT: Mark, Melanie R.
; APPLICANT: Scadden, David T.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Baron, Will F.
; TITLE OF INVENTION: Protein Tyrosine Kinases
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080

COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/445,640
FILING DATE: 22-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/170558
FILING DATE: 20-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/157563
FILING DATE: 23-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Hasak, Janet E.
REGISTRATION NUMBER: 28,616
REFERENCE/DOCKET NUMBER: 854C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1896
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 913 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-445-640-4

Query Match 100.0%; Score 2175; DB 1; Length 913;
Best Local Similarity 100.0%; Pred No. 3.3e-216;
Matches 399; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DADMKGHFDPKCRYALGMQDRTPDSDISASSSSWSDSTAARHSRLSSDGDGAWCPAGS 60
Db 19 DADMKGHFDPKCRYALGMQDRTPDSDISASSSSWSDSTAARHSRLSSDGDGAWCPAGS 78
QY 61 VFPEEYLYQVLDLQRLHLVALVGTQGRHAGGLGKFSRSYRLRYSDGRRWGKDRWQ 120
Db 79 VFPEEYLYQVLDLQRLHLVALVGTQGRHAGGLGKFSRSYRLRYSDGRRWGKDRWQ 138

QY 121 EVISGNEDEPGVVLKDLGPPMVARLVREYPRADRVMSVCLRVLYGCLWRDGLLSYAPV 180
DB 139 EVISGNEDEPGVVLKDLGPPMVARLVREYPRADRVMSVCLRVLYGCLWRDGLLSYAPV 198
QY 181 GQTYLSEAVYLNDSYDGHVTGGVGLGQLADGVVGLDDFRKSOELRWPFYDYGW 240
DB 199 GQTYLSEAVYLNDSYDGHVTGGVGLGQLADGVVGLDDFRKSOELRWPFYDYGW 258
QY 241 SNHFSFGVYVMEFEFDRLRAFAQMVHCNNMHTLGARLPGGVCECRPRRGPAMAWEGEPM 300
DB 259 SNHFSFGVYVMEFEFDRLRAFAQMVHCNNMHTLGARLPGGVCECRPRRGPAMAWEGEPM 318
QY 301 RHNLGNLGDPRARAVSVPLGGRVARELQCRFLFAGPWLLESEISFISDVVNNSSPALGG 360
DB 319 RHNLGNLGDPRARAVSVPLGGRVARELQCRFLFAGPWLLESEISFISDVVNNSSPALGG 378
QY 361 TFPAPMPPPPPTNFSSLELEPRGQOPVAKPEGSPTA 399
DB 379 TFPAPMPPPPPTNFSSLELEPRGQOPVAKPEGSPTA 417

RESULT 2

US-08-445-640-8
; Sequence 8, Application US/08445640
; Patent No. 5709858

GENERAL INFORMATION:

APPLICANT: Godowski, Paul J.
APPLICANT: Mark, Melanie R.
APPLICANT: Scadden, David T.
APPLICANT: Baker, Kevin P.
APPLICANT: Baron, Will F.
TITLE OF INVENTION: Protein Tyrosine Kinases
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 5.25 inch, 360 kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/445,640
FILING DATE: 22-MAY-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/170558
FILING DATE: 20-DEC-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/157563
FILING DATE: 23-NOV-1993

ATTORNEY/AGENT INFORMATION:

NAME: Hasak, Janet E.
REGISTRATION NUMBER: 28,616
REFERENCE/DOCKET NUMBER: 854C2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415/225-1896
TELEFAX: 415/952-9881
TELEX: 910/371-7168

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:
LENGTH: 399 amino acids
TYPE: amino acid

TOPOLOGY: linear

US-08-445-640-8

Query Match 100.0%; Score 2175; DB 1; Length 399;
Best Local Similarity 100.0%; Pred. No. 1e-216;

Matches 399; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DADMKGHFDPAKCRVALGMDQRTIPDSISASSSSWSDSTAARHSRLESDDGDGAWCPAGS 60
DB 1 DADMKGHFDPAKCRVALGMDQRTIPDSISASSSSWSDSTAARHSRLESDDGDGAWCPAGS 60
QY 61 VFPKEEYLVQDLQRLHLVALVGTGRRHAGGLGKEFSRSYRLRYSRDGRRWGKDRWGQ 120
DB 61 VFPKEEYLVQDLQRLHLVALVGTGRRHAGGLGKEFSRSYRLRYSRDGRRWGKDRWGQ 120
QY 121 EVISGNEDEPGVVLKDLGPPMVARLVREYPRADRVMSVCLRVLYGCLWRDGLLSYAPV 180
DB 121 EVISGNEDEPGVVLKDLGPPMVARLVREYPRADRVMSVCLRVLYGCLWRDGLLSYAPV 180
QY 181 GQTYLSEAVYLNDSYDGHVTGGVGLGQLADGVVGLDDFRKSOELRWPFYDYGW 240
DB 181 GQTYLSEAVYLNDSYDGHVTGGVGLGQLADGVVGLDDFRKSOELRWPFYDYGW 240
QY 241 SNHFSFGVYVMEFEFDRLRAFAQMVHCNNMHTLGARLPGGVCECRPRRGPAMAWEGEPM 300
DB 241 SNHFSFGVYVMEFEFDRLRAFAQMVHCNNMHTLGARLPGGVCECRPRRGPAMAWEGEPM 300
QY 301 RHNLGNLGDPRARAVSVPLGGRVARELQCRFLFAGPWLLESEISFISDVVNNSSPALGG 360
DB 301 RHNLGNLGDPRARAVSVPLGGRVARELQCRFLFAGPWLLESEISFISDVVNNSSPALGG 360
QY 361 TFPAPMPPPPPTNFSSLELEPRGQOPVAKPEGSPTA 399
DB 361 TFPAPMPPPPPTNFSSLELEPRGQOPVAKPEGSPTA 399

RESULT 3

US-08-336-343A-2

; Sequence 2, Application US/08336343A
; Patent No. 5677144

GENERAL INFORMATION:

APPLICANT: Ullrich, Axel
APPLICANT: Alves, Frauke
TITLE OF INVENTION: CCK-2, A No. 5677144el Receptor Tyrosine Kinase
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/336,343A
FILING DATE: 08-NOV-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7683-065

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090
TELEFAX: (212) 790-9090
TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 919 amino acids
TYPE: amino acid

TOPOLOGY: unknown

MOLECULE TYPE: protein

US-08-336-343A-2

Query Match 99.6%; Score 2167; DB 1; Length 919;
Best Local Similarity 99.7%; Pred. No. 2.2e-215;
Matches 398; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DADMKGHFPAKCRYALGMDRTIPDSISASSSSDSTAARHSRLSSDGDGACWPAGS 60
DB 19 DADMKGHFPAKCRYALGMDRTIPDSISASSSSDSTAARHSRLSSDGDGACWPAGS 78
QY 61 VPKBEYEQVLDLQRLHVALVGTQGRHAGGLGKEFSRSYRLRYSRDGRRWGWKDRWGQ 120
DB 79 VPKBEYEQVLDLQRLHVALVGTQGRHAGGLGKEFSRSYRLRYSRDGRRWGWKDRWGQ 138
QY 121 EVISGNEDEGVVLDKGLPMPVARLVRFYPRADRVMSVCLRVLYGCLWRDGLLSYTA 180
DB 139 EVISGNEDEGVVLDKGLPMPVARLVRFYPRADRVMSVCLRVLYGCLWRDGLLSYTA 198
QY 181 GQTMYLEAVYLNDSYDGTGGVGLQGLADGVGLDFFRKSQELRVWPGDYDVGW 240
DB 199 GQTMYLEAVYLNDSYDGTGGVGLQGLADGVGLDFFRKSQELRVWPGDYDVGW 258
QY 241 SNHSFSGYVEMEFEDRLRAFOAMQVHCNNMHTLGARLPGGVCECRFRRGPAWAGEP 300
DB 259 SNHSFSGYVEMEFEDRLRAFOAMQVHCNNMHTLGARLPGGVCECRFRRGPAWAGEP 318
QY 301 RNLGNLGDPRARAVSVPLGGVRFVLCRFLFAGPWLLEISFISDVVNSSPALGG 360
DB 319 RNLGNLGDPRARAVSVPLGGVRFVLCRFLFAGPWLLEISFISDVVNSSPALGG 378
QY 361 TFPAPWPPGPPPTNFSSLELEPRGQQPVAKPEGSP 399
DB 379 TFPAPWPPGPPPTNFSSLELEPRGQQPVAKPEGSP 417

RESULT 4

US-08-336-343A-4
; Sequence 4, Application US/08336343A
; Patent No. 5677144
; GENERAL INFORMATION:
; APPLICANT: Ullrich, Axel
; APPLICANT: Alives, Frauke
; TITLE OF INVENTION: CCK-2, A No. 5677144el Receptor Tyrosine Kinase
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/336,343A
; FILING DATE: 08-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7683-065
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 855 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-336-343A-4

Query Match 46.1%; Score 1002.5; DB 1; Length 855;
Best Local Similarity 53.6%; Pred. No. 4.2e-95;
Matches 196; Conservative 54; Mismatches 109; Indels 7; Gaps 6;

QY 5 KGHFPAKCRYALGMDRTIPDSISASSSSDSTAARHSRLSSDGDGACWPAGSVFPK 64
DB 22 KAQYNPAICRYPLGMSGQIPDEDITASSQWSESTAARGLDSEEDGACWCEIPVEPD 81
QY 65 E-EYLVQDLQRLHVALVGTQGRHAGGLGKEFSRSYRLRYSRDGRRWGWKDRWGQ 123
DB 82 DLKFLQDLHVALVGTQGRHAGGLGKEFSRSYRLRYSRDGRRWGWKDRWGQ 141
QY 124 SGNEDPEGVVLKDLGPPMPVARLVRFYPRADRVMSVCLRVLYGCLWRDGLLSYTA 183
DB 142 DGSNPDYDIFLKDLEPPIVARFVRFIPVTDHSMNVCNRVLYGCVLWDLGLVSNAPAG 201
QY 184 MYL--SEAVYLNDSYDGTGGVGLQGLADGVGLDFFRKSQELRVWPGDYDVGW 241
DB 202 FVLPGGSIILYLNDSYDGTGGVGLQGLADGVGLDFFRKSQELRVWPGDYDVGW 260
QY 242 NHFSFGYVEMEFEDRLRAFOAMQVHCNNMHTLGARLPGGVCECRFRRGPAWAGEP 301
DB 261 NESATNGYIEMFEDRLRAFOAMQVHCNNMHTLGARLPGGVCECRFRRGPAWAGEP 319
QY 302 HNLGNLGDPRARAVSVPLGGVRFVLCRFLFAGPWLLEISFISDVVNSSPALGG 360
DB 320 FVLVDDVNPASRVTVPLHHRMASAIKCOYHEADTWMESEITFQSDAANYNSEAL-P 378
QY 361 TFPAP 366
DB 379 TSPMAP 384

RESULT 5

US-08-456-647B-20
; Sequence 20, Application US/08456647B
; Patent No. 5811516
; GENERAL INFORMATION:
; APPLICANT: Lenke Ph.D. et al., Greg E.
; TITLE OF INVENTION: PROTEIN-TYROSINE KINASE GENES
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/456,647B
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/237,401
; FILING DATE: 02-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/884,486
; FILING DATE: 15-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Wetherell Ph.D., John R.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: 07251/007002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 678-5070
; TELEFAX: (619) 678-5099
; INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:
LENGTH: 854 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-456-647B-20

Query Match 45.8%; Score 995.5; DB 2; Length 854;
Best Local Similarity 49.6%; Pred. No. 2.2e-94;
Matches 198; Conservative 58; Mismatches 116; Indels 27; Gaps 7;

QY 5 KGHDPKACRYALGMDRTIPDSISASSWSDSSTAARHSRLSSDGDGAWCPAGSVFPK 64
DB 22 KAQVNPACRYPLGMSGGHIPDEDITASSQWSESTAARHRLDSEEGDGCWCPPIPVQPD 81
QY 65 E-EYQLVDLQRLHLVALVGTQGRHAGLGKEFSRSLRYSDGRWGMKDRWGQEV 123
DB 82 DLKEFLQIDLRLTHFTILVGTQGRHAGHGIEFAPMKINYSRDSRWISWRNRHGQVL 141
QY 124 SGNEDPEGVVLKDLGPPMVARLVRFYPRADRVMSVCLRVLYGCLWRDGLLSYAPVQOT 183
DB 142 DGSNPYDVFLKLEPPIVARFVRLIPVTDHSMNVMCMRVLYGCVWLDGLVSNAPAGQ 201
QY 184 MYL--SEAVYLDSTYDGHVTGGYGLGQLADGVVGLDDFRKSOELRWVPGYDYVGS 241
DB 202 FVLPGGSIYLDNSVYDG-AVGYSMTEGLGQLTDGVSLDDFTQTHEYHVMVPGYDYVGR 260
QY 242 NHFSSGVMEFEFDRLRFAQMVHNCNNHMTLGARLPGGVCEFRFRGPAMAVEGPMR 301
DB 261 NESATNGFIEMFEFDRIRNFTTKVHCNMFAGVKIFKEVQCYF-RSEASEWEPTAVY 319
QY 302 HNLGNLGDPRARAVSVPLGGRVARFLQCRFLFAGPWLLEFSEISFISD--VYNNSSPALG 359
DB 320 FPLVLDVNPASAREVTVPLHHRMASAIKQVHFADTWMFSEITFQSDAAMYNNS----- 375
QY 360 GTFPAPWPPGPPPTNFSSLELEPRGQPVAKPEGSP 398
DB 375 GALPTSP-----MAPTYDPMKVDSDNT 398

RESULT 6

US-08-237-401A-20
Sequence 20, Application US/08237401A
Patent No. 5837448
GENERAL INFORMATION:
APPLICANT: Lemke Ph.D. et al., Greg E.
TITLE OF INVENTION: PROTEIN-TYROSINE KINASE GENES
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/237,401A
FILING DATE: 02-MAY-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/684,486
FILING DATE: 15-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Haile Ph.D., Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07251/007001
TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 678-5070
TELEFAX: (619) 678-5099
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 854 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-237-401A-20

Query Match 45.8%; Score 995.5; DB 2; Length 854;
Best Local Similarity 49.6%; Pred. No. 2.2e-94;
Matches 198; Conservative 58; Mismatches 116; Indels 27; Gaps 7;

QY 5 KGHDPKACRYALGMDRTIPDSISASSWSDSSTAARHSRLSSDGDGAWCPAGSVFPK 64
DB 22 KAQVNPACRYPLGMSGGHIPDEDITASSQWSESTAARHRLDSEEGDGCWCPPIPVQPD 81
QY 65 E-EYQLVDLQRLHLVALVGTQGRHAGLGKEFSRSLRYSDGRWGMKDRWGQEV 123
DB 82 DLKEFLQIDLRLTHFTILVGTQGRHAGHGIEFAPMKINYSRDSRWISWRNRHGQVL 141
QY 124 SGNEDPEGVVLKDLGPPMVARLVRFYPRADRVMSVCLRVLYGCLWRDGLLSYAPVQOT 183
DB 142 DGSNPYDVFLKLEPPIVARFVRLIPVTDHSMNVMCMRVLYGCVWLDGLVSNAPAGQ 201
QY 184 MYL--SEAVYLDSTYDGHVTGGYGLGQLADGVVGLDDFRKSOELRWVPGYDYVGS 241
DB 202 FVLPGGSIYLDNSVYDG-AVGYSMTEGLGQLTDGVSLDDFTQTHEYHVMVPGYDYVGR 260
QY 242 NHFSSGVMEFEFDRLRFAQMVHNCNNHMTLGARLPGGVCEFRFRGPAMAVEGPMR 301
DB 261 NESATNGFIEMFEFDRIRNFTTKVHCNMFAGVKIFKEVQCYF-RSEASEWEPTAVY 319
QY 302 HNLGNLGDPRARAVSVPLGGRVARFLQCRFLFAGPWLLEFSEISFISD--VYNNSSPALG 359
DB 320 FPLVLDVNPASAREVTVPLHHRMASAIKQVHFADTWMFSEITFQSDAAMYNNS----- 375
QY 360 GTFPAPWPPGPPPTNFSSLELEPRGQPVAKPEGSP 398
DB 375 GALPTSP-----MAPTYDPMKVDSDNT 398

RESULT 7

US-08-670-707A-37
Sequence 37, Application US/08670707A
Patent No. 5859204
GENERAL INFORMATION:
APPLICANT: Lollar, John S.
TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: USA
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/670,707A
FILING DATE: 26-JUN-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US94/13200
FILING DATE: 15-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/212,133

FILING DATE: 11-MAR-1994
PRIOR APPLICATION DATA: US 07/864,004
APPLICATION NUMBER: 07-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Greenlee, Lorance L.
REGISTRATION NUMBER: 27,894
REFERENCE/DOCKET NUMBER: 75-95F
TELECOMMUNICATION INFORMATION:
TELEPHONE: 303/499-8080
TELEFAX: 303/499-8089
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 2133 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-670-707A-37

Query Match 9.7%; Score 210; DB 2; Length 2133;
Best Local Similarity 34.7%; Pred. No. 1.2e-12;
Matches 59; Conservative 24; Mismatches 71; Indels 16; Gaps 5;
QY 9 DPAKCRYALGMQDRTIPDSISASS-----SWSDDTAARHSRLSSDGDGAWCPAGSV 61
Db 1971 DLNCSMPLGMQNAISDSQITASSLSNIFATWSPQA-----RLHLOGRTNARPRVS- 2026
QY 62 FPKEEYLQVLDLQRLHLVALVGTQGRHAGGLGKEFSRSLRYSDGRRWGKDRWGQE 121
Db 2026 --SAEEMQLQVLDLQRLHLVALVGTQGRHAGGLGKEFSRSLRYSDGRRWGKDRWGQE 121
QY 122 VISGNEDEPGVVLKDLGPPVAVLRVRYPRADRVMSVCLRVLYGCLWRD 171
Db 2083 VFOGNDSTPTVVALDPPFLTRYLRHP-TSWAQHIALRLVLCGEAOD 2131

RESULT 8
US-08-670-707A-39
Sequence 39, Application US/08670707A
Patent No. 5859204
GENERAL INFORMATION:
APPLICANT: Lollar, John S.
TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: USA
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/670,707A
FILING DATE: 26-JUN-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US94/13200
FILING DATE: 15-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/212,133
FILING DATE: 11-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/864,004
FILING DATE: 07-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Greenlee, Lorance L.
REGISTRATION NUMBER: 27,894

REFERENCE/DOCKET NUMBER: 75-95F
TELECOMMUNICATION INFORMATION:
TELEPHONE: 303/499-8080
TELEFAX: 303/499-8089
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 1443 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-670-707A-39

Query Match 9.7%; Score 210; DB 2; Length 1443;
Best Local Similarity 34.7%; Pred. No. 6.8e-13;
Matches 59; Conservative 24; Mismatches 71; Indels 16; Gaps 5;
QY 9 DPAKCRYALGMQDRTIPDSISASS-----SWSDDTAARHSRLSSDGDGAWCPAGSV 61
Db 1281 DLNCSMPLGMQNAISDSQITASSLSNIFATWSPQA-----RLHLOGRTNARPRVS- 1336
QY 62 FPKEEYLQVLDLQRLHLVALVGTQGRHAGGLGKEFSRSLRYSDGRRWGKDRWGQE 121
Db 1336 --SAEEMQLQVLDLQRLHLVALVGTQGRHAGGLGKEFSRSLRYSDGRRWGKDRWGQE 121
QY 122 VISGNEDEPGVVLKDLGPPVAVLRVRYPRADRVMSVCLRVLYGCLWRD 171
Db 1393 VFOGNDSTPTVVALDPPFLTRYLRHP-TSWAQHIALRLVLCGEAOD 1441

RESULT 9
US-08-111-939-2
Sequence 2, Application US/08111939
Patent No. 5460951
GENERAL INFORMATION:
APPLICANT: Kawai, Shinji
APPLICANT: Takeshita, Sunao
APPLICANT: Okazaki, Makoto
APPLICANT: Amann, Egon
TITLE OF INVENTION: Bone-Related Carboxypeptidase-Like
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESS: Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/111,939
FILING DATE: 26-AUG-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 324033/92
FILING DATE: 03-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 230029/92
FILING DATE: 28-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: Forman, David S.
REGISTRATION NUMBER: 33,694
REFERENCE/DOCKET NUMBER: 02481.1321-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4000
INFORMATION FOR SEQ ID NO: 2:

; ANTI-SENSE: NO

NET REVENUE/DOCKET: CRT 00 000

Query Match 8.8%; Score 192; DB 1; Length 2351;
Best Local Similarity 31.8%; Pred. No. 1e-10;
Matches 54; Conservative 32; Mismatches 68; Indels 16; Gaps 5;

Qy 9 DPAAKRYALGMODRTIPDSDISASS-----SNWSDTAARHSLRLESSPDGCAWPAGSV 61
| |||| : ||| : ||| : ||| : ||| :
| |||| : ||| : ||| : ||| : ||| :

```

Query Watch      8.8%; Score 192; DB 1; Length 2332;
Best Local Similarity 31.8%; Pred. No. 9.8e-11;
Matches 54; Conservative 32; Mismatches 68; Indels 16; Gaps

QY 9 DPACYRYALGMQDRTIPDSISASS-----SWSDSTAARHSRLSSDGDGACNCPAGSV 61
      | : | | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 2170 DLNCSMPLGMESKATSDAQITASSYFTNMFATWSPSKARLHIGRSN----AMRPQVN- 2225

QY 62 FPKREBYLQVDLQRHLVALVCTQGRAGGLKEFSRYRLYSRDGRRWGWKDRNGQE 121
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 2225 --NPKRWLQVDFQKTKMKTGVITGTVGKS--LITSMYVKBFLLSSSDGQHWITLFFONGKVK 2281

```


GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 1999, 22:52:09 ; Search time 22.34 Seconds
(without alignments)
423.042 Million cell updates/sec

Title: US-08-170-558-8
Perfect score: 2175
Sequence: 1 DADMKGHFPAKCRYALGMQ.....LELEPRGQVPVAKPEGSPTA 399

Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues

Database: A_Geneseq_36.*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2167	99.6	919	1	R75502	Human mammary carc
2	2167	99.6	919	1	R75504	Human mammary carc
3	2167	99.6	919	1	W34672	Human mammary carc
4	2167	99.6	882	1	W34673	Human mammary carc
5	2167	99.6	876	1	W34675	Human mammary carc
6	2124.5	97.7	914	1	R71100	Protein-tyrosine-k
7	1002.5	46.1	855	1	R75503	Human colonic aden
8	1002.5	46.1	855	1	R75505	Human colonic aden
9	1002.5	46.1	855	1	W34671	CCR-2, a human mam
10	1002.5	46.1	855	1	W77114	Discoidin domain r
11	995.5	45.8	854	1	W79152	Receptor protein t
12	995.5	45.8	854	1	W81409	Receptor protein t
13	425	19.5	650	1	R54089	Partial sequence o
14	364	16.7	563	1	W34674	Human mammary carc
15	222	10.2	909	1	W96254	Mouse semaphorin r
16	222	10.2	926	1	W96255	Human semaphorin r
17	220	10.1	909	1	W96249	Rat semaphorin rec
18	220	10.1	909	1	W96250	Mouse semaphorin r
19	220	10.1	914	1	W96251	Mouse semaphorin r
20	220	10.1	926	1	W96252	Mouse semaphorin r
21	220	10.1	931	1	W96253	Mouse semaphorin r
22	220	10.1	901	1	W96256	Mouse semaphorin r
23	220	10.1	906	1	W96257	Mouse semaphorin r
24	217	10.0	925	1	W96308	Neuropilin-2. New
25	211.5	9.7	845	1	W36816	Human E2A-binding
26	210	9.7	1443	1	W44137	Human E2A-binding
27	210	9.7	2133	1	W44133	Sus scrofa factor
28	204	9.4	923	1	W96246	Human semaphorin r
29	204	9.4	923	1	W96248	Mouse semaphorin r
30	201.5	9.3	1128	1	R49994	Mouse carboxypepti
31	201.5	9.3	1128	1	W36817	Mouse E2A-binding
32	199	9.1	921	1	W96247	Rat semaphorin rec
33	199	9.1	921	1	W96309	Neuropilin. New is
34	194	8.9	768	1	W36823	Human E2A-binding
35	193	8.9	2351	1	P70448	Human Factor VIII:
36	193	8.9	2351	1	R78223	Human Factor VIII:
37	193	8.9	768	1	W36815	Human E2A-binding
38	192	8.8	1424	1	P91169	Sequence of 740 Ar
39	192	8.8	1424	1	P80268	Modified factor VI
40	192	8.8	1516	1	P80265	Modified factor VI
41	192	8.8	1425	1	P80267	Modified factor VI
42	192	8.8	2351	1	P81113	Factor VIII encode
43	192	8.8	2351	1	P80659	Sequence of human

ALIGNMENTS

RESULT 1

R75502
ID R75502 standard; Protein: 919 AA.
AC R75502;
DE 26-NOV-1995 (first entry)
KW Human mammary carcinoma kinase 10 (MCK-10).
KW Mammary carcinoma kinase 10; transmembrane receptor;
KW receptor tyrosine kinase; cancer.
OS Homo sapiens.
FH Key
FT Location/Qualifiers
FT peptide 1..18
FT /label= signal
FT domain 31..185
FT /label= discoidin I-like domain
FT cleavage_site 304..307
FT /label= putative precursor cleavage site
FT region 417..439
FT /label= transmembrane
FT misc_difference 505..541
FT /label= alternatively spliced sequence I
FT misc_difference 666..671
FT /label= alternatively spliced sequence II
FT misc_difference 25..42
FT /label= NT alpha
FT /note= "peptide antibody recognition site"
FT misc_difference 309..321
FT /label= NT beta
FT /note= "see above"
FT misc_difference 909..919
FT /label= CT beta
FT /note= "see above"
WO9514088-A.
26-MAY-1995.
16-NOV-1994; E03797.
16-NOV-1993; US-153397.
(PLAC) MAX PLANCK GES. FOERDERUNG WISSENSCHAFTEN.
Alves FHE, Ullrich A;
WPI; 95-224054/29.
N-PSDB; Q92520.
New nucleic acid encoding MCK-10 receptor tyrosine kinase - and
derived vectors, transformed cells, proteins and antibodies useful
for diagnosis and treatment of proliferative disease, esp. cancer,
and for screening modulators
disclosure; Page 53-55; 115pp; English.
CDNA prep. from human breast cancer cell line MCF7 (ATCC HTB22) and
used in a PCR with two degenerate oligo primer pools based on
conserved sequences of the kinase domain of receptor tyrosine
kinases. One clone, designated MCK-10, was identified as novel RTK.
The PCR fragment was used to screen a lambda gtl1 library of human
fetal brain cDNA. Several overlapping clones were identified. The
composite of these cDNA clones is given in Q92520 and the deduced AA
sequence in R75502. Some of the clones had a deletion of 6 AAs at
posn. 2315 in the MCK-10 sequence. MCK-10 has all the
characteristics of a receptor TK (see R75502 FT). Screening of
human placental library yielded two cDNA clones MCK-10-1 and
MCK-10-2. One of the clones isolated from the human fetal brain
library contd. an additional 18 nts in the TK domain. The MCK-10 splice
isoforms have been designated MCK-10-1 (with an additional 11 bp between
nts 1832 and 1943); MCK-10-2 (without any insertions); MCK-10-3 (with the
additional 11 bp and and 18 bp in the TK domain); and MCK-10-4 (with the
proreceptors are 101.13 and 97.17 kD respectively, and can thus be
subdivided into a 34.31 kD alpha subunit and and 66.84 or 62.88 kD
beta subunits that contain the TK homology and alternative splice sites.
Sequence 919 AA;

Factor VIII:c varia
Human factor VIII

	PT	New nucleic acid encoding CK-2 receptor tyrosine kinase - and derived vectors, transformed cells, proteins and antibodies, useful for diagnosis and treatment of proliferative and nervous system diseases and for screening modulators
	PS	Disclosure: Page 70-72; 115pp; English.
	CC	cDNA prep'd. from human breast cancer cell line MCF7 (ATCC HTB22); was used in a PCR with two degenerate oligo primer pools based on conserved sequences of the kinase domain of receptor tyrosine kinases. One clone, designated MCK-10, was identified as novel RTK. The PCR fragment was used to screen a lambda gtlI library of human fetal brain cDNA. Several overlapping clones were identified. The composite of these cDNA clones is given in Q92522 and the deduced AA sequence in R75504. Some of the clones had a deletion of 6AA at posn. 2315 in the MCK-10 sequence. MCK-10 has all the characteristics of a receptor PK (see R75504 Ff). Screening of human placental library yielded two cDNA clones. One of the clones isolated from the human fetal brain library contained an additional 18 nts in the TK domain. The MCK-10 splice isoforms have been designated MCK-10-1 (with an additional 111 bp between nts 1832 and 1943); MCK-10-2 (without any insertions); MCK-10-3 (with the additional 111 bps and 18 bp in the TK domain); and MCK-10-4 (with the additional 111 bps and The predicted mol. wts. of MCK-10-1 and MCK-10-2 proreceptors are 101.13 and 97.17 kD respectively, and can thus be subdivided into a 34.31 kD alpha subunit and a 66.84 or 62.88 kD beta subunits that contain the TK homology and alternative splice sites.
	SQ	Sequence 919 AA;
		Query Match 99.6%; Score 2167; DB 1; Length 919; Best Local Similarity 99.7%; Pred. No. 7.8e-210; Matches 398; Conservative 0; Mismatches 1; Indels 0; Gaps
	QY	1 DADMGHFDPAKRYALGMDRIPDSISASSWSDDTAARHSLESDDGDGAWCPAGS 60
	Db	19 DADMGHFDPAKRYALGMDRIPDSISASSWSDDTAARHSLESDDGDGAWCPAGS 78
	QY	61 VPKKEEYLQVDLQRHLVALVTQGRHAGGLGKFBSRYLRYSRDGRRRMWGWKDRWGQ 120
	Db	79 VPKKEEYLQVDLQRHLVALVTQGRHAGGLGKFBSRYLRYSRDGRRRMWGWKDRWGQ 138
	QY	121 EVISGNEDPEGVLKDLGPPMARVRLREYPADRVMVSLCLRVLYGCLRWDGLLSYTAPV 180
	Db	139 EVISGNEDPEGVLKDLGPPMARVRLVFPRADRVMSVLCRLVELYGCLRWDGLLSYTAPV 198
	QY	181 GOTMYLSEAVYLNDSYDGHTVGGLGYGGLGADGVVGLDDFRKSQELRWPGDYDVGW 240
	Db	199 GOTMYLSEAVYLNDSYDGHTVGGLGYGGLGADGVVGLDDFRKSQELRWPGDYDVGW 258
	QY	241 SNHSPSSCYGVMEFEFDRLEAFAMQVHCNNMHTLGARLPGVCRFRGPANAWEGEPM 300
	Db	259 SNHSPSSCYGVMEFEFDRLEAFAMQVHCNNMHTLGARLPGVCRFRGPANAWEGEPM 318
	QY	301 RHNLGNLGDPRARAVSVPLGGVRAREFLQCRLEFAGPMLLFSEISFTSDVVNNSPALGG 360
	Db	319 RHNLGNLGDPRARAVSVPLGGVRAREFLQCRLEFAGPMLLFSEISFTSDVVNNSPALGG 378
	QY	361 TFPAPWWPGPPPTNFSSLELEPRGQQPVAKPEGSPTA 399
	Db	379 TFPAPWWPGPPPTNFSSLELEPRGQQPVAKAEGSTA 417
	RESULT	3
	W34672	
	ID	- W34672 standard; Protein: 919 AA.
	AC	W34672:
	DT	17-FEB-1998 (first entry)
	DE	Human mammary carcinoma kinase 10 (MCK-10) amino acid sequence.
	KW	Mammary carcinoma kinase; MCK-10; receptor tyrosine kinase.
	KW	Proliferative disease; cancer; insulin receptor family;
	KW	Tyrosine kinase neurotrophin receptor; MCK-10 activity;
	KW	neurological disorder; aberrant expression.
	OS	Homo sapiens.
	FH	Key Location/Qualifiers

FT Peptide 1. .18
 FT /label= signal_sequence
 FT Protein 19. .919
 FT /note= "mature_protein"
 FT Domain 31. .185
 FT /label= Discoidin_I_like_domain
 FT Cleavage_site 304. .307
 FT /label= endopeptidase_furin
 FT /note= "putative precursor cleavage site"
 FT Region 48. .439
 FT Binding_site 617. .627
 FT /label= ATP_binding_motif
 FT Modified_site 802. .803
 FT /label= autophosphorylation_sites
 FT /note= "putative"
 FT Modified_site 798
 FT /label= autophosphorylation_site
 FT /note= "putative"
 FT Binding_site 844. .847
 FT /label= binding_motif_for_P13_kinase
 FT /note= "binding motif for phosphatidylinositol 3'
 FT kinase"
 FT Binding_site 832. .832
 FT /label= potential_substrate_binding_site
 FT Binding_site 506. .509
 FT /label= putative_receptor_binding_site_for_SHC
 FT /note= "SHC is an oncogenic SH2 domain containing
 FT molecule"
 FT Binding_site 510. .513
 FT /label= GTPase_activity_protein_binding_site
 FT /note= "putative"
 FT Region 505. .541
 FT /note= "alternatively spliced variant"
 FT Region 666. .671
 FT /note= "alternatively spliced sequence"
 FT Region 26. .42
 FT /note= "antibody recognition sequence NTalpha"
 FT Region 309. .321
 FT /note= "antibody recognition sequence NTbeta"
 FT Region 902. .919
 FT /note= "antibody recognition sequence CTbeta"
 FT US5677144-A.
 PN 14-OCT-1997.
 PD 08-NOV-1994; 336343.
 PR 16-NOV-1993; US-153397.
 PA (ALVE/) ALVES F H E.
 PA (ULLR/) ULLRICH A.
 PI Alves FHE, Ullrich A;
 DR WPI; 97-511869/47.
 DR N-PSDB; T93785.
 PT Truncated receptor tyrosine kinase CCK-2 - and nucleic acid coding
 PT for it, useful for cancer diagnosis
 PS Disclosure; Fig 1; 70pp; English.
 CC The present sequence represents the protein sequence of a mammary
 CC carcinoma kinase, called MCK-10. This kinase belongs to a novel family
 CC of receptor tyrosine kinases, and expression is associated with
 CC proliferative diseases such as cancer. The MCK-10 receptor tyrosine
 CC kinase has extensive sequence similarity to the insulin receptor family.
 CC The MCK-10 gene was isolated by PCR using 2 degenerate oligonucleotide
 CC primer pools, using a template cDNA synthesised by reverse transcription
 CC of poly-A RNA from the human mammary carcinoma cell line MCF7. The MCK-10
 CC protein contains 2 alternative spliced sequences, from amino acids
 CC 505-541 and 666-671. The sequence represented by amino acids 585-595 may
 CC be important, as deletion of this motif in the activin receptor
 CC serine/threonine kinase results in reduced ligand binding affinity.
 CC MCK-10 is expressed in brain tissue, and the protein shares homology with
 CC the tyrosine kinase neurotrophin receptor. Modulation of MCK-10 activity
 CC therefore may be used for treatment of neurological disorders. MCK-10 is
 CC also expressed in a variety of cancer cell lines and tumour tissue. The
 CC nucleotide sequence of MCK-10, or parts of it, can be used for diagnostic
 CC purposes to detect aberrant expression of MCK-10 genes. Inhibitors of
 CC MCK-10 receptor activity may have therapeutic value in the treatment of

CC diseases such as cancer.
 SQ Sequence 919 AA;
 Query Match 99.6%; Score 2167; DB 1; Length 919;
 Best Local Similarity 99.7%; Pred. No. 7.8e-210;
 Matches 398; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 DADMKGHFDPKCRYALGMDRTIPDSDISASSWSNSTAARHSRLSSDGDGAWCPAGS 60
 Db 19 DADMKGHFDPKCRYALGMDRTIPDSDISASSWSNSTAARHSRLSSDGDGAWCPAGS 78
 Qy 61 VFPEEYLYVDLQRLHLVALVGTQGRHAGLGKFEKFSYRLRYSRDGRWGMKDRWGQ 120
 Db 79 VFPEEYLYVDLQRLHLVALVGTQGRHAGLGKFEKFSYRLRYSRDGRWGMKDRWGQ 138
 Qy 121 EVISGNEDEPGVVLKDLGPPMVARLVRFYPRADRVMSVCLRVLYGLWRDGLLSYAPV 180
 Db 139 EVISGNEDEPGVVLKDLGPPMVARLVRFYPRADRVMSVCLRVLYGLWRDGLLSYAPV 198
 Qy 181 GQTYLSEAVYLNDSYDGHVTGGLQYGGGLQGLADGVVGLDDFKKSOELRWPMGYDYVGW 240
 Db 199 GQTYLSEAVYLNDSYDGHVTGGLQYGGGLQGLADGVVGLDDFKKSOELRWPMGYDYVGW 258
 Qy 241 SNHSFSSGYVEMEFEDRLRAFOAMQVHCNNMHTLGARLPGGVECRFRPGPAMAWEGEPM 300
 Db 259 SNHSFSSGYVEMEFEDRLRAFOAMQVHCNNMHTLGARLPGGVECRFRPGPAMAWEGEPM 318
 Qy 301 RHNLGNGLDPRARAVSVPLGGVRVRFLOCRFLFAGPWLFLFSEISFISDVVNNSSPALGG 360
 Db 319 RHNLGNGLDPRARAVSVPLGGVRVRFLOCRFLFAGPWLFLFSEISFISDVVNNSSPALGG 378
 Qy 361 TFPAPWMPGPPPTNFSSLEPRGQOPVAKPEGSPTA 399
 Db 379 TFPAPWMPGPPPTNFSSLEPRGQOPVAKPEGSPTA 417
 RESULT 4
 W34673
 ID W34673 standard; Protein; 882 AA.
 AC W34673;
 DT 17-FEB-1998 (first entry)
 DE Human mammary carcinoma kinase 10 (MCK-10) splice variant 1.
 KW Mammary carcinoma kinase; MCK-10; receptor tyrosine kinase;
 KW proliferative disease; cancer; insulin receptor family;
 KW tyrosine kinase neurotrophin receptor; MCK-10 activity;
 KW neurological disorder; aberrant expression.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Peptide 1. .18
 FT /label= signal_sequence
 FT Protein 19. .919
 FT /note= "mature_protein"
 FT Domain 31. .185
 FT /label= Discoidin_I_like_domain
 FT Cleavage_site 304. .307
 FT /label= endopeptidase_furin
 FT /note= "putative precursor cleavage site"
 FT Region 48. .439
 FT /label= transmembrane_region
 FT Binding_site 580. .590
 FT /label= ATP_binding_motif
 FT Modified_site 765. .766
 FT /label= autophosphorylation_sites
 FT /note= "putative"
 FT Modified_site 761
 FT /label= autophosphorylation_site
 FT /note= "putative"
 FT Binding_site 807. .810
 FT /label= binding_motif_for_P13_kinase
 FT /note= "binding motif for phosphatidylinositol 3'
 FT kinase"
 FT Binding_site 795. .795

FT Region /label= potential_substrate_binding_site
 FT 26..42 /note= "antibody recognition sequence Ntalpha"
 FT 309..321 /note= "antibody recognition sequence Ntbeta"
 FT 865..882 /note= "antibody recognition sequence Ctbeta"
 FT US5677144-A.
 PN 14-OCT-1997.
 PD 08-NOV-1994; 336343.
 PR 16-NOV-1993; US-153397.
 PA (ALVE/) ALVES F H E.
 PA (ULLR/) ULLRICH A.
 PI Alves FHE, Ullrich A;
 DR WPI; 97-511869/47.
 PT Truncated receptor tyrosine kinase CKK-2 - and nucleic acid coding
 PT for it, useful for cancer diagnosis
 PS Disclosure: Page -: 70pp: English.
 CC The present sequence represents a splice variant of a mammary
 CC carcinoma kinase (MCK-10). This kinase belongs to a novel family
 CC of receptor tyrosine kinases, and expression is associated with
 CC proliferative diseases such as cancer. The MCK-10 receptor tyrosine
 CC kinase has extensive sequence similarity to the insulin receptor family.
 CC The MCK-10 gene was isolated by PCR using 2 degenerate oligonucleotide
 CC primer pools, using a template cDNA synthesised by reverse transcription
 CC of poly-A RNA from the human mammary carcinoma cell line MCF7. The
 CC amplified PCR product was used to screen human foetal brain and placental
 CC libraries, from which the present splice variant was isolated. This
 CC splice variant does not possess amino acids 505-541 of MCK-10 (W34672).
 CC The sequence represented by amino acids 548-558 may be important, as
 CC deletion of this motif in the activin receptor serine/threonine kinase
 CC results in reduced ligand binding affinity. MCK-10 is expressed in brain
 CC tissue, and the protein shares homology with the tyrosine kinase
 CC neurotrophin receptor. Modulation of MCK-10 activity therefore may be used
 CC for treatment of neurological disorders. MCK-10 is also expressed in a
 CC variety of cancer cell lines and tumour tissue. The nucleotide sequence
 CC of MCK-10, or parts of it, can be used for diagnostic purposes to detect
 CC aberrant expression of MCK-10 genes. Inhibitors of MCK-10 (or splice
 CC variants) receptor activity may have therapeutic value in the treatment
 CC of diseases such as cancer.
 CC note: the present sequence does not appear in the specification, but was
 CC created using information provided.
 SQ Sequence 882 AA;

Query Match 99.6%; Score 2167; DB 1; Length 882;
 Best Local Similarity 99.7%; Pred. No. 7.4e-210;
 Matches 398; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DADMKGFDPKCRYALGMDRTIPDSISASSWSUSTAAHRSRLSSDGDGAWCPAGS 60
 DB 19 DADMKGFDPKCRYALGMDRTIPDSISASSWSUSTAAHRSRLSSDGDGAWCPAGS 78
 QY 61 VFPKEEYLQVDLQRLHLVALVGTQGRHAGLGKFEFSYRLYSRDRRWGKDRWGQ 120
 DB 79 VFPKEEYLQVDLQRLHLVALVGTQGRHAGLGKFEFSYRLYSRDRRWGKDRWGQ 138
 QY 121 EYISGNEDEGVVVKDGLPPMVARLVRYFPRADRVMSVCLRVLYGCLWRDGLLSYAPV 180
 DB 139 EYISGNEDEGVVVKDGLPPMVARLVRYFPRADRVMSVCLRVLYGCLWRDGLLSYAPV 198
 QY 181 GOTMYLSEAVYLNDSYDGHVTGGVGGYGLGQLADGVVGLDDFRKSELRWPGDYVGV 240
 DB 199 GOTMYLSEAVYLNDSYDGHVTGGVGGYGLGQLADGVVGLDDFRKSELRWPGDYVGV 258
 QY 241 SNHFSFGYVEMEFEDRLRAFAQMVHCHNNHTLGLARLPGGVCECRFRGPAMAWEGEPM 300
 DB 259 SNHFSFGYVEMEFEDRLRAFAQMVHCHNNHTLGLARLPGGVCECRFRGPAMAWEGEPM 318
 QY 301 RHNLGNGLDPRARAVSVPLGGRVARELQCRFLFAGPWLIFSEISFISDVYNNSSPALGG 360
 DB 319 RHNLGNGLDPRARAVSVPLGGRVARELQCRFLFAGPWLIFSEISFISDVYNNSSPALGG 378

QY 361 TFPAPWMPGPPPTNFSSLELEPRGQPVAKPEGSPTA 399
 DB 379 TFPAPWMPGPPPTNFSSLELEPRGQPVAKPEGSPTA 417
 RESULT 5
 W34675
 ID W34675 standard; Protein: 876 AA.
 AC W34675;
 DT 17-FEB-1998 (first entry)
 DE Human mammary carcinoma kinase 10 (MCK-10) splice variant 3.
 KW Mammary carcinoma kinase; MCK-10; receptor tyrosine kinase;
 KW proliferative disease; cancer; insulin receptor family;
 KW tyrosine kinase neurotrophin receptor; MCK-10 activity;
 KW neurological disorder; aberrant expression.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Peptide 1..118
 FT /label= signal_sequence
 FT 19..876
 FT /note= "mature_protein"
 FT Domain 31..185
 FT /label= Discoidin_I_like_domain
 FT Cleavage_site 304..307
 FT /label= endopeptidase_furin
 FT /note= "putative precursor cleavage site"
 FT Region 48..439
 FT /label= transmembrane_region
 FT Binding_site 580..590
 FT /label= ATP_binding_motif
 FT Modified_site 760..761
 FT /label= autophosphorylation_sites
 FT /note= "putative"
 FT Modified_site 756..756
 FT /label= autophosphorylation_site
 FT /note= "putative"
 FT Binding_site 802..805
 FT /label= binding_motif_for_P13_kinase
 FT /note= "binding motif for phosphatidylinositol 3'
 FT kinase"
 FT Binding_site 790
 FT /label= potential_substrate_binding_site
 FT 26..42
 FT /note= "antibody recognition sequence Ntalpha"
 FT Region 309..321
 FT /note= "antibody recognition sequence Ntbeta"
 FT Region 860..877
 FT /note= "antibody recognition sequence Ctbeta"
 FT US5677144-A.
 PN 14-OCT-1997.
 PD 08-NOV-1994; 336343.
 PR 16-NOV-1993; US-153397.
 PA (ALVE/) ALVES F H E.
 PA (ULLR/) ULLRICH A.
 PI Alves FHE, Ullrich A;
 DR WPI; 97-511869/47.
 PT Truncated receptor tyrosine kinase CKK-2 - and nucleic acid coding
 PT for it, useful for cancer diagnosis
 PS Disclosure: Page -: 70pp: English.
 CC The present sequence represents a splice variant of a mammary
 CC carcinoma kinase (MCK-10). This kinase belongs to a novel family
 CC of receptor tyrosine kinases, and expression is associated with
 CC proliferative diseases such as cancer. The MCK-10 receptor tyrosine
 CC kinase has extensive sequence similarity to the insulin receptor family.
 CC The MCK-10 gene was isolated by PCR using 2 degenerate oligonucleotide
 CC primer pools, using a template cDNA synthesised by reverse transcription
 CC of poly-A RNA from the human mammary carcinoma cell line MCF7. The
 CC amplified PCR product was used to screen human foetal brain and placental
 CC libraries, from which the present splice variant was isolated. This
 CC splice variant does not possess amino acids 505-541 or 666-671 of MCK-10
 CC (W34672). The sequence represented by amino acids 548-558 may be
 CC important, as deletion of this motif in the activin receptor
 CC serine/threonine kinase results in reduced ligand binding affinity.

CC MCK-10 is expressed in brain tissue, and the protein shares homology with
 CC the tyrosine kinase neurotrophin receptor. Modulation of MCK-10 activity
 CC therefore may be used for treatment of neurological disorders. MCK-10 is
 CC also expressed in a variety of cancer cell lines and tumour tissue. The
 CC nucleotide sequence of MCK-10, or parts of it, can be used for diagnostic
 CC purposes to detect aberrant expression of MCK-10 genes. Inhibitors of
 CC MCK-10 (or splice variants) receptor activity may have therapeutic value
 CC in the treatment of diseases such as cancer.
 CC note: the present sequence does not appear in the specification, but was
 CC created using information provided.
 SQ Sequence 876 AA;

Query Match 99.6%; Score 2167; DB 1; Length 876;
 Best Local Similarity 99.7%; Pred. No. 7.3e-210;
 Matches 398; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DADMKGHFDPACRYALGMDRTIPDSISASSWSSTAAHRSRLSSDGDGAWCPAGS 60
 DB 19 DADMKGHFDPACRYALGMDRTIPDSISASSWSSTAAHRSRLSSDGDGAWCPAGS 78
 QY 61 VFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKEFSRSYRLYSRDGRMMWGWDRWQ 120
 DB 79 VFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKEFSRSYRLYSRDGRMMWGWDRWQ 138
 QY 121 EVISGNEDEPGVWLKDLGPPMVARLVRFYPRADRVMSVCLRVLYGCLWRDGLLSYAPV 180
 DB 139 EVISGNEDEPGVWLKDLGPPMVARLVRFYPRADRVMSVCLRVLYGCLWRDGLLSYAPV 198
 QY 181 GQWYLSAAYLNDSTYDGTGVLGGLQGLADGVVGLDDPRKSOELRVWPGDYVGV 240
 DB 199 GQWYLSAAYLNDSTYDGTGVLGGLQGLADGVVGLDDPRKSOELRVWPGDYVGV 258
 QY 241 SNHSFSSGYVEMEFEDRLRAFOAMOVHCNMHTLGARLPGGVCEFRFRGPAMAWEGEPM 300
 DB 259 SNHSFSSGYVEMEFEDRLRAFOAMOVHCNMHTLGARLPGGVCEFRFRGPAMAWEGEPM 318
 QY 301 RHNGLGNLGDPRARAVSVPLGGRVAREFLQCRFLFAGPWLIFSEISFISDVVNSSPALGG 360
 DB 319 RHNGLGNLGDPRARAVSVPLGGRVAREFLQCRFLFAGPWLIFSEISFISDVVNSSPALGG 378
 QY 361 TFPAPWPPGPPPTNFSSLELEPRGQPVAKPEGSPTA 399
 DB 379 TFPAPWPPGPPPTNFSSLELEPRGQPVAKAEGSPTA 417

RESULT 6
 R71100
 ID R71100 standard; Protein; 914 AA.
 AC R71100;
 DT 17-AUG-1995 (first entry)
 DE Protein-tyrosine-kinase PTK22.
 KW Protein-tyrosine-kinase; PTK; discoidin domain receptor; cancer;
 KW breast tumor; mamma carcinoma; diagnosis; prognosis; therapy.
 OS Homo sapiens.
 PN WO9502187-A.
 PD 19-JAN-1995.
 PF 08-JUL-1994; G01480.
 PR 09-JUL-1993; GB-014271.
 PA (CANC-) CANCER RES INST.
 PA (WELL) WELLCOME FOUND LTD.
 PI Barker KT, Crompton MR, Gusterson BA, Martindale JE;
 PI Mitchell PJ, Page MJ, Spence P;
 DR WPI: 95-066991/09.
 DR N-PSDB: Q84782.
 PT Method for screening substances, using protein tyrosine kinase -
 PT for potential utility as therapeutic agents for cancer
 PS Disclosure; Page 26-30; 51pp; English.
 CC cDNA derived from tumor metastatic tissue was amplified using
 CC primers (given in Q84783-84) based on sequences (R71101, R71103)
 CC associated with protein-tyrosine-kinases (PTK). Novel PTK22 was
 CC identified in an isolated subclone. The 3' sequence of PTK22 was
 CC obtained by reverse transcription (using the primer of Q84786) and

CC PCR amplification (primers Q84787-88) of RNA of human breast
 CC carcinoma cell line MDA MG 468. The partial DNA sequence of PTK22
 CC is given in Q84782.
 SQ Sequence 914 AA;

Query Match 97.7%; Score 2124.5; DB 1; Length 914;
 Best Local Similarity 98.2%; Pred. No. 1.5e-205;
 Matches 393; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

QY 1 DADMKGHFDPACRYALGMDRTIPDSISASSWSSTAAHRSRLSSDGDGAWCPAGS 60
 DB 19 DADMKGHFDPACRYALGMDRTIPDSISASSWSSTAAHRSRLSSDGDGAWCPAGS 78
 QY 61 VFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKEFSRSYRLYSRDGRMMWGWDRWQ 120
 DB 79 VFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKEFSRSYRLYSRDGRMMWGWDRWQ 138
 QY 121 EVISGNEDEPGVWLKDLGPPMVARLVRFYPRADRVMSVCLRVLYGCLWRDGLLSYAPV 180
 DB 139 EVISGNEDEPGVWLKDLGPPMVARLVRFYPRADRVMSVCLRVLYGCLWRDGLLSYAPV 198
 QY 181 GQWYLSAAYLNDSTYDGTGVLGGLQGLADGVVGLDDPRKSOELRVWPGDYVGV 240
 DB 199 GQWYLSAAYLNDSTYDGTGVLGGLQGLADGVVGLDDPRKSOELRVWPGDYVGV 258
 QY 241 SNHSFSSGYVEMEFEDRLRAFOAMOVHCNMHTLGARLPGGVCEFRFRGPAMAWEGEPM 300
 DB 259 SNHSFSSGYVEMEFEDRLRAFOAMOVHCNMHTLGARLPGGVCEFRFRGPAMAWEGEPM 318
 QY 301 RHNGLGNLGDPRARAVSVPLGGRVAREFLQCRFLFAGPWLIFSEISFISDVVNSSPALGG 360
 DB 319 RHNGLGNLGDPRARAVSVPLGGRVAREFLQCRFLFAGPWLIFSEISFISDVVNSSPALGG 378
 QY 361 TFPAPWPPGPPPTNFSSLELEPRGQPVAKPEGSPTA 399
 DB 379 TFPAPWPPGPPPTNFSSLELEPRGQPVAKAEGSPTA 418

RESULT 7
 R75503
 ID R75503 standard; Protein; 855 AA.
 AC R75503;
 DT 26-NOV-1995 (first entry)
 DE Human colonic adenocarcinoma kinase 2 (CCK-2).
 KW Mammary carcinoma kinase 10; MCK-10; transmembrane receptor; CCK-2;
 KW receptor tyrosine kinase; colonic adenocarcinoma kinase 2; cancer.
 OS Homo sapiens.
 PN WO9514088-A.
 PD 26-MAY-1995.
 PF 16-NOV-1994; E03797.
 PR 16-NOV-1993; US-153397.
 PA (PLAC) MAX PLANKK GES FOERDERUNG WISSENSCHAFTEN.
 PI Alves FHE, Ullrich A;
 DR WPI: 95-224054/29.
 DR N-PSDB: Q92521.
 PT New nucleic acid encoding MCK-10 receptor tyrosine kinase - and
 PT derived vectors, transformed cells, proteins and antibodies useful
 PT for diagnosis and treatment of proliferative disease, esp. cancer,
 PT and for screening modulators
 PS Disclosure; Page 57-60; 115pp; English.
 CC A member of the mammary carcinoma kinase 10 (MCK-10) receptor
 CC tyrosine kinase family was identified using a PCR (with two degenerate
 CC oligo primer pools based on conserved sequences of the kinase domain of
 CC receptor tyrosine kinases) and cDNA prep. from colonic
 CC adenocarcinoma RNA. The nt sequence of the novel receptor,
 CC designated CCK-2, is given in Q92521 and the deduced AA sequence in
 CC R75503. Analysis of the CCK-2 nt and AA sequence indicated
 CC significant homology with MCK-10 throughout the extracellular,
 CC transmembrane and intracellular regions. The regions of homology
 CC extend into the N-terminus consensus sequence for the discoidin I
 CC like family of proteins.
 SQ Sequence 855 AA;

FT Region 400..421 binding site"
FT /label= transmembrane_region
FT /note= "putative"
FT 30...185
FT /label= Discoidin_I_like_domain
FT 433..438
FT /label= protein_kinase_C_binding_site
FT /note= "putative"
PN US5677144-A.
PD 14-OCT-1997.
PF 08-NOV-1994; 336343.
PR 16-NOV-1993; US-153397.
PA (ALVE/) ALVES F H E.
PA (ULLR/) ULLRICH A.
PI Alves FHE, Ullrich A;
DR WPI; 97-511869/47.
DR N-PSDB; T93784.
PT Truncated receptor tyrosine kinase CCK-2 - and nucleic acid coding
PT for it, useful for cancer diagnosis
PS Claim 5; Fig 3; 70pp; English.
CC The present sequence represents the amino acid sequence of human CCK-2, a
CC member of the mammary carcinoma kinase 10 (MCK-10, W34672) family of
CC receptor tyrosine kinases. The protein contains a remarkably high
CC number of proline residues arranged as PXXP or PXXR repeats, suggesting
CC a random coil structure for the hydrophilic juxtamembrane region. This
CC region is probably a major domain for interactions with cellular
CC substrates and other regulatory proteins. Expression of CCK-2 is
CC associated with proliferative diseases such as cancer. The CCK-2 gene was
CC identified by PCR and a cDNA prepared from colonic adenocarcinoma RNA.
CC CCK-2 is expressed in a wide variety of cancer cell lines and tumour
CC tissue. The CCK-2 nucleic acids can be used for diagnostic purposes to
CC detect aberrant expression of CCK-2 genes. Engineered cell lines,
CC containing recombinant vectors with the present sequence, are useful for
CC producing infectious retroviral particles. The cell lines may also be
CC used to evaluate and screen drugs involved in CCK-2 activation and
CC regulation.
SQ Sequence 855 AA;

Query Match 46.1%; Score 1002.5; DB 1; Length 855;
Best Local Similarity 53.6%; Pred. No. 1.8e-92;
Matches 196; Conservative 54; Mismatches 109; Indels 7; Gaps 6;
QY 5 KGHDPKACRYALGMQDRTIPDSISASSWSSTAAHRSLESSDGDGAWCPAGSVFPPK 64
DB 22 KAQVNPALCRPLGMSGGQIPDEDITASSQWSESTAKYGRDSEEDGAWCPPIPEPD 81
QY 65 E-EYLVQDQLRHLVALVGTQGRHAGLGKGFERSYRLRYSRDGRMMGWKDRWGOEVI 123
DB 82 DLKEFLQIDLHTLFTLVGTQGRHAGHGIEFAPWKYKINYSRDGTWISWRNRHGKQVL 141
QY 124 SGNEDPGVVLKDLGPPMVARLRFYPRADRYMSVCLRVLYGCLWRDGLLSYTPVGT 183
DB 142 DGSNPNYDIFLKDLEPPIVARFVRFPVTDHSMNVCNRVELYGCWLDGLVSYNAPAGQ 201
QY 184 MYL--SEAVYLNDSYDGTGGVGLQYGLGOLADGVVGLDDFRKSOELRVWPGDYVWGS 241
DB 202 FVLPGGSIYLNDSYDG-AVGYSWTEGLGQLTDGVSLDDFTQTHEYHWWPGDYVWGR 260
QY 242 NHSSSGVMEFEFDRLRAFQAMQVHCNNMHTIGARLPGGVCEFRFRGPMAMWEGEPMR 301
DB 261 NESATNGYIEIMFEDRIRNFTTKVHCNNMFAKGVKIFKEVQCYF-RSEASEWEPNAIS 319
QY 302 HNLGNLGDPRARAVSVPLGGRVARFLQCRFLFAGPMLLSEISFTSD-VYNNSSPALGG 360
DB 320 FPLVLDVNPASAREVTVPVPLHHRMASAIKCOYHFADTWMFSEITFQSDAAMYNSEAL-P 378
QY 361 TFPAP 366
DB 379 TSPMAP 384

RESULT 11

W79152

ID W79152 standard; Protein; 854 AA.

AC W79152;

DT 19-NOV-1998 (first entry)

DE Receptor protein tyrosine kinase (PTK) subtype tyro-10.

KW PTK; receptor; protein tyrosine kinase; brain tissue.

RESULT 10
W77114
ID W77114 standard; Protein; 855 AA.
AC W77114; 1998 (first entry)
DT 16-NOV-1998 (first entry)
DE Discoidin domain receptor 2 protein.
KW Discoidin domain receptor; transformation; metastasis; collagen; ss;
KW Cleidocranial dysplasia; Sicker syndrome; extracellular matrix; MMP-1.
OS Homo sapiens.
PN W09834954-A2.
PD 13-AUG-1998.
PF 05-FEB-1998; CA0093.
PR 06-FEB-1997; US-041578.
PA (MOUN) MOUNT SINAI HOSPITAL CORP.
PI Pawson A, Vogel W;
DR WPI; 98-447168/38.
DR N-PSDB; V48292.
PT Novel ligands of discoidin domain receptor tyrosine kinase,
PT especially collagen - useful for treating e.g. metastasis,
PT cleidocranial dysplasia or Sicker syndrome
PS Disclosure: Fig 22a; 115pp; English.
CC The discoidin domain receptor (DDR) can be used to identify and evaluate
CC substances which affect DDR receptor tyrosine kinase signalling pathways
CC in the cell. Compounds which modulate such signalling pathways can be
CC used to alter transformation or metastasis in mammals, to treat
CC conditions involving structural or functional deregulation of collagens,
CC e.g. Cleidocranial dysplasia or Sicker syndrome, conditions requiring
CC modulation of extracellular matrix synthesis, degradation or remodelling,
CC or to treat conditions needing modulation of MMP-1 expression such as
CC wound healing.
SQ Sequence 855 AA;

Query Match 46.1%; Score 1002.5; DB 1; Length 855;
Best Local Similarity 53.6%; Pred. No. 1.8e-92;
Matches 196; Conservative 54; Mismatches 109; Indels 7; Gaps 6;

QY 5 KGHDPKACRYALGMQDRTIPDSISASSWSSTAAHRSLESSDGDGAWCPAGSVFPPK 64
DB 22 KAQVNPALCRPLGMSGGQIPDEDITASSQWSESTAKYGRDSEEDGAWCPPIPEPD 81
QY 65 E-EYLVQDQLRHLVALVGTQGRHAGLGKGFERSYRLRYSRDGRMMGWKDRWGOEVI 123
DB 82 DLKEFLQIDLHTLFTLVGTQGRHAGHGIEFAPWKYKINYSRDGTWISWRNRHGKQVL 141
QY 124 SGNEDPGVVLKDLGPPMVARLRFYPRADRYMSVCLRVLYGCLWRDGLLSYTPVGT 183
DB 142 DGSNPNYDIFLKDLEPPIVARFVRFPVTDHSMNVCNRVELYGCWLDGLVSYNAPAGQ 201
QY 184 MYL--SEAVYLNDSYDGTGGVGLQYGLGOLADGVVGLDDFRKSOELRVWPGDYVWGS 241
DB 202 FVLPGGSIYLNDSYDG-AVGYSWTEGLGQLTDGVSLDDFTQTHEYHWWPGDYVWGR 260
QY 242 NHSSSGVMEFEFDRLRAFQAMQVHCNNMHTIGARLPGGVCEFRFRGPMAMWEGEPMR 301
DB 261 NESATNGYIEIMFEDRIRNFTTKVHCNNMFAKGVKIFKEVQCYF-RSEASEWEPNAIS 319
QY 302 HNLGNLGDPRARAVSVPLGGRVARFLQCRFLFAGPMLLSEISFTSD-VYNNSSPALGG 360
DB 320 FPLVLDVNPASAREVTVPVPLHHRMASAIKCOYHFADTWMFSEITFQSDAAMYNSEAL-P 378
QY 361 TFPAP 366
DB 379 TSPMAP 384


```

FT modified_site 167 /note= "Potential N-linked glycosylation site."
FT domain 195..216 /label= Transmembrane domain.
FT region 365..370 /label= ATP binding region.
FT /note= "Highly conserved among protein tyrosine
FT kinase enzymes."
FT
PN DE4239817-A.
PD 01-JUN-1994.
PD 26-NOV-1992; 239817.
PF 26-NOV-1992; DE-239817.
PR (CHEM-) CHEMOTHERAPEUTISCHES FORSCHUNG.
PI Holtrich U, Ruebsamen-waigmann H, Strebhardt K;
DR WPI; 94-184380/23.
DR N-PSDB; 064158.
PT New protein tyrosine kinase and related nucleic acid - vectors,
PT transformed cells, etc., useful for diagnosis and treatment of
PT tumours
PS Claim 1; Page 8; 9pp; German.
CC The gene encoding this polypeptide is related to the trk
CC proto-oncogene. Antibodies against the polypeptide are useful for
CC diagnosis and for the treatment of tumours. The antibodies may also
CC be radiolabelled or coupled to a cytotoxin for destruction of cancer
CC cells. Antisense nucleic acid can be used to inhibit gene
CC expression.
CC Sequence 650 AA;
SQ

Query Match 19.5%; Score 425; DB 1; Length 650;
Best Local Similarity 50.6%; Pred. No. 2e-34;
Matches 90; Conservative 26; Mismatches 58; Indels 4; Gaps 4;

QY 190 VYLNDSYDGHYVGGYGLGGLADGVGLDDFRKSOELRWPGYDYVGNHHSFSSGY 249
DB :||||| ||| || ||||| ||| ||||| ::| ||||| ||| ||| ::||
5 IYLNDSYVDG-AVGYSWTEGLGQLTDGVSLGLDDFTQTHEYVWPGYDYVGNHHSFSSGY 63
QY 250 VEEFEFDRLRAFMQVQVHNNHMTLGARLPGGVCECRFRGPPAMAVEGPMRHLNGLG 309
DB ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
64 IEINFEEDRLRNTTMMKVCNNHFAKGVKFEVQCYF-RSEASEWPAISFFPLVDV 122
QY 310 DPRARVSVPLGGRVARFTQCRFLRAGPWLIFSEISFISD-VYNNSSPALGGTFFPPAP 366
DB :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
123 NPSARFVTVPLHHRMASAIKQYHFADTWMMFSEITFQSDAAMYNSEAL-PTSPMAP 179

RESULT 14
W34674
ID W34674 standard; Protein; 563 AA.
AC W34674;
DT 17-FEB-1998 (first entry)
DE Human mammary carcinoma kinase 10 (MCK-10) splice variant 2..
KW Mammary carcinoma kinase; MCK-10; receptor tyrosine kinase;
KW Proliferative disease; cancer; insulin receptor family;
KW tyrosine kinase neurotrophin receptor; MCK-10 activity;
KW neurological disorder; aberrant expression.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Peptide 1..18
FT /label= signal_sequence
FT Protein 19..919
FT /note= "mature_protein"
FT Domain 31..185
FT /label= Discoicidin_I_like_domain
FT Cleavage_site 304..307
FT /label= endopeptidase_furin
FT /note= "putative precursor cleavage site"
FT Region 48..439
FT /label= transmembrane_region
FT Binding_site 617..627
FT /label= ATP_binding_motif
FT Modified_site 797..798
FT /label= autophosphorylation_sites

```

```

FT Modified_site 793 /note= "putative"
FT /label= autophosphorylation_site
FT /note= "putative"
FT Binding_site 839..842
FT /label= binding_motif_for_P13_kinase
FT /note= "binding motif for phosphatidylinositol 3'
FT kinase"
FT Binding_site 827..827
FT /label= potential_substrate_binding_site
FT Binding_site 506..509
FT /label= putative_receptor_binding_site_for_SHC
FT /note= "SHC is an oncogenic SH2 domain containing
FT molecule"
FT Binding_site 510..513
FT /label= GTPase_activity_protein_binding_site
FT /note= "putative"
FT Region 26..42
FT /note= "antibody recognition sequence NTalpha"
FT Region 309..321
FT /note= "antibody recognition sequence NTbeta"
FT Region 897..913
FT /note= "antibody recognition sequence CTbeta"
FT
PN US5677144-A.
PD 14-OCT-1997.
PD 08-NOV-1994; 336343.
PF 16-NOV-1993; US-153397.
PR (ALVE/) ALVES F H E.
PA (ULLR/) ULLRICH A.
PI Alves FHE, Ullrich A;
DR WPI; 97-511869/47.
PT Truncated receptor tyrosine kinase CCK-2 - and nucleic acid coding
PT for it, useful for cancer diagnosis
PT Disclosure; Page -: 70pp; English.
CC The present sequence represents a splice variant of a mammary
CC carcinoma kinase (MCK-10). This kinase belongs to a novel family
CC of receptor tyrosine kinases, and expression is associated with
CC proliferative diseases such as cancer. The MCK-10 receptor tyrosine
CC kinase has extensive sequence similarity to the insulin receptor family.
CC The MCK-10 gene was isolated by PCR using 2 degenerate oligonucleotide
CC primer pools, using a template cDNA synthesised by reverse transcription
CC of poly-A RNA from the human mammary carcinoma cell line MCF7. The
CC amplified PCR product was used to screen human foetal brain and placental
CC libraries, from which the present splice variant was isolated. This
CC splice variant does not possess amino acids 666-671 of MCK-10 (W34672).
CC The sequence represented by amino acids 585-595 may be important, as
CC deletion of this motif in the activin receptor serine/threonine kinase
CC results in reduced ligand binding affinity. MCK-10 is expressed in brain
CC tissue, and the protein shares homology with the tyrosine kinase
CC neurotrophin receptor. Modulation of MCK-10 activity therefore may be used
CC for treatment of neurological disorders. MCK-10 is also expressed in a
CC variety of cancer cell lines and tumour tissue. The nucleotide sequence
CC of MCK-10, or parts of it, can be used for diagnostic purposes to detect
CC aberrant expression of MCK-10 genes. Inhibitors of MCK-10 (or splice
CC variants) receptor activity may have therapeutic value in the treatment
CC of diseases such as cancer.
CC note: the present sequence does not appear in the specification, but was
CC created using information provided.
SQ Sequence 563 AA;

```

```

Query Match 16.7%; Score 364; DB 1; Length 563;
Best Local Similarity 98.5%; Pred. No. 2.4e-28;
Matches 66; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 333 LFAGPWLFFSEISFISDVNNSSPALGGTFFPPAPWPGPPPTNFSSLELEPRGQOPVAK 392
DB ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
1 LFAGPWLFFSEISFISDVNNSSPALGGTFFPPAPWPGPPPTNFSSLELEPRGQOPVAK 60
QY 393 PEGSPTA 399
DB |||||
61 AEGSPTA 67

```


GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 4, 1999, 05:43:21 ; Search time 1266.11 seconds
(without alignments)
9135.687 Million cell updates/sec

Title: US-08-170-558-3
Perfect score: 3637
Sequence: 1 GAATTCGAGTCGACGTTG.....AAAAAAAAAAAAAAAAAAAA 3637

Scoring table: IDENTITY_NUC

Searched: 679419 seqs, 1590154680 residues

Database : GenEmbl.*

- 1: gb_ba1.*
- 2: gb_ba2.*
- 3: gb_om.*
- 4: gb_ov.*
- 5: gb_pat.*
- 6: gb_ph.*
- 7: gb_pl1.*
- 8: gb_pl2.*
- 9: gb_pr1.*
- 10: gb_pr2.*
- 11: gb_pr3.*
- 12: gb_ro.*
- 13: gb_st.*
- 14: gb_sts.*
- 15: gb_sy.*
- 16: gb_un.*
- 17: gb_vi.*
- 18: em_fun.*
- 19: em_hg.*
- 20: em_hum1.*
- 21: em_hum2.*
- 22: em_in.*
- 23: em_om.*
- 24: em_or.*
- 25: em_ov.*
- 26: em_pat.*
- 27: em_ph.*
- 28: em_pl.*
- 29: em_ro.*
- 30: em_sts.*
- 31: em_sy.*
- 32: em_un.*
- 33: em_vi.*
- 34: gb_hgt1.*
- 35: gb_hgt2.*
- 36: gb_hgt3.*
- 37: gb_in2.*
- 38: gb_ba1.*
- 39: gb_ba2.*
- 40: em_hum3.*
- 41: em_hum4.*
- 42: gb_pr4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	3637	100.0	3637	5	I80845	I80845 Sequence 3

2	3451	94.9	3962	5	I68738
3	3441.8	94.6	3754	5	A42378
4	3438.6	94.5	3738	9	HUMRTK
5	3399.6	93.5	3803	9	HUMCAK
6	3243.6	89.2	3554	10	HSTRKE
7	3232.4	88.9	3841	10	HSRETYK1
8	3167.6	87.1	3692	10	HUMCAKA
9	2375.8	65.3	3659	12	AF026259
10	2354.8	64.7	3674	12	MUSCAK
11	2314	63.6	3743	12	RATPTK3D
12	1197	32.9	1197	5	I80847
c	837.2	23.0	48731	11	AC004211
13	826.6	22.7	1072	10	HSDRTK17
14	803.6	22.1	12010	11	HSU48705
15	669.4	18.4	3259	12	MMTYRO10
16	666.8	18.3	3157	5	I68739
17	666.8	18.3	3157	5	I68740
c	666.8	18.3	3096	10	HSRPTK
19	570.4	15.7	660	12	S77585
20	340	9.3	549	12	S77556
21	326.6	9.0	845	10	HSDRTK6
22	250.2	6.9	970	10	HSDRTK12
23	250.2	6.9	970	10	HSDRTK12
24	250	6.9	250	14	G15457
c	243	6.7	355	10	HSDRTK15
26	241.4	6.6	768	12	AF016247
27	236.2	6.5	1229	10	HSDRTK123
28	226.4	6.2	661	10	HSDRTK14
29	218	6.0	400	14	G11190
30	200	5.5	2738	4	GGU43396
31	198.4	5.5	1568	10	HSTRK2H
32	196.8	5.4	2727	4	GGTRKA
33	196.8	5.4	2010	10	HSTRKT3ON
34	193.6	5.3	2301	5	I96186
35	193.6	5.3	2701	9	HUMTRKPOA
36	193.6	5.3	2301	10	HSTRKR
37	193.6	5.3	1746	10	HSTRKT1
38	190	5.2	3707	5	I35538
39	190	5.2	3194	10	S76473
40	190	5.2	3707	11	HSU12140
41	185.8	5.1	2697	12	RATTRKCA
42	185.8	5.1	2682	12	RATTRKCN3
43	185.8	5.1	2697	12	S62924
44	185.2	5.1	1408	11	AF052184
45	184.4	5.1	2067	36	GCTK

ALIGNMENTS

RESULT	1
I80845	
LOCUS	I80845
DEFINITION	Sequence 3 from patent US 5709858.
ACCESSION	I80845
NID	93209135
VERSION	I80845.1
KEYWORDS	GI:3209135
SOURCE	Unknown.
ORGANISM	Unknown.
REFERENCE	1 (bases 1 to 3637)
AUTHORS	Godowski, P.J., Mark, M.R. and Scadden, D.T.
TITLE	Antibodies specific for Rse receptor protein tyrosine kinase
JOURNAL	Patent: US 5709858-A 3 20-JAN-1998;
FEATURES	Location/Qualifiers
source	1..3637
BASE COUNT	721 a 1094 c 1079 g 743 t
ORIGIN	

Query Match 100.0%; Score 3637; DB 5; Length 3637;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 3637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY 1	GAATTCGAGTCGAGCTTGAAGAAATGCAAGAGATGCTGCCACACCCCTT	60	
Db 1	GAATTCGAGTCGAGCTTGAAGAAATGCAAGAGATGCTGCCACACCCCTT	60	
QY 61	AGGCCCCGAGGATCAGGAGCTATGGACACAGAGCCCTGTCACTTTACTGCTGCTC	120	
Db 61	AGGCCCCGAGGATCAGGAGCTATGGACACAGAGCCCTGTCACTTTACTGCTGCTC	120	
QY 121	TTGGTGGCAAGTGGAGATGCTCACATGAAGGACATTTTGATCTCGCAAGTCCGCTAT	180	
Db 121	TTGGTGGCAAGTGGAGATGCTCACATGAAGGACATTTTGATCTCGCAAGTCCGCTAT	180	
QY 181	GCCTGGGCATCAGAGCCGACCATCCAGACAGTACATCTCTGCTTCAGCTCCTGG	240	
Db 181	GCCTGGGCATCAGAGCCGACCATCCAGACAGTACATCTCTGCTTCAGCTCCTGG	240	
QY 241	TCAGATTCACACTGCGCCCGCCACAGCAGGTTGGAGACAGTACAGGAGTGGGGCTGG	300	
Db 241	TCAGATTCACACTGCGCCCGCCACAGCAGGTTGGAGACAGTACAGGAGTGGGGCTGG	300	
QY 301	TGCCCCGAGGTCGCTGTTCCCAAGCAGGAGGAGTACTTGCAGGTGGATCTACAACGA	360	
Db 301	TGCCCCGAGGTCGCTGTTCCCAAGCAGGAGGAGTACTTGCAGGTGGATCTACAACGA	360	
QY 361	CTGACCTGCTGGTCTGGTGGGACACGAGGACGGCATGCCGGGGCTGGGCAAGGAG	420	
Db 361	CTGACCTGCTGGTCTGGTGGGACACGAGGACGGCATGCCGGGGCTGGGCAAGGAG	420	
QY 421	TTCTCCCGAGCTACCGGCTGCGTACTTCCCGGATGGTCCCGCTGGATGGGCTGAAG	480	
Db 421	TTCTCCCGAGCTACCGGCTGCGTACTTCCCGGATGGTCCCGCTGGATGGGCTGAAG	480	
QY 481	GACCGCTGGGTGAGGAGTATCTAGGCAATGAGGACCTGAGGAGTGGTCTCAAG	540	
Db 481	GACCGCTGGGTGAGGAGTATCTAGGCAATGAGGACCTGAGGAGTGGTCTCAAG	540	
QY 541	GACCTGGGCCCCCGATGGTGGGACACGAGTGGTCCGCTTACCCCCGGGTGACCGGGTC	600	
Db 541	GACCTGGGCCCCCGATGGTGGGACACGAGTGGTCCGCTTACCCCCGGGTGACCGGGTC	600	
QY 601	ATGAGCGCTCTGTCGGGTGAGCTCTATGGCTGCTCTGGAGGATGACTCTGCT	660	
Db 601	ATGAGCGCTCTGTCGGGTGAGCTCTATGGCTGCTCTGGAGGATGACTCTGCT	660	
QY 661	TACACCGCCCCGTGGGACACAAATGATTTATCTCAGGCGCTGTACCTCAAGACTCC	720	
Db 661	TACACCGCCCCGTGGGACACAAATGATTTATCTCAGGCGCTGTACCTCAAGACTCC	720	
QY 721	ACCTATGACGACATACCGTGGGCGGACTCGAGTATGGGGTCTGGCCAGCTGGCAGAT	780	
Db 721	ACCTATGACGACATACCGTGGGCGGACTCGAGTATGGGGTCTGGCCAGCTGGCAGAT	780	
QY 781	GCTGTGTGGGCTGGATGACTTTAGGAAGTACAGAGCTGGGGTCTGGCCAGGCTAT	840	
Db 781	GCTGTGTGGGCTGGATGACTTTAGGAAGTACAGAGCTGGGGTCTGGCCAGGCTAT	840	
QY 841	GACTATGTGGGATGGACCAACACAGCTTCTCCAGTGGCTATGTGGAGATGGAATTGAG	900	
Db 841	GACTATGTGGGATGGACCAACACAGCTTCTCCAGTGGCTATGTGGAGATGGAATTGAG	900	
QY 901	TTTCACCGGCTGAGGGCTTCCAGGCTATCGAGTCCAGTGTAAACATGACACGCTG	960	
Db 901	TTTCACCGGCTGAGGGCTTCCAGGCTATCGAGTCCAGTGTAAACATGACACGCTG	960	
QY 961	GGAGCCGCTGCTGCGGGGGTGAATGTCGCTTCCGGGCTGGCCCTGCTATGGGCTG	1020	
Db 961	GGAGCCGCTGCTGCGGGGGTGAATGTCGCTTCCGGGCTGGCCCTGCTATGGGCTG	1020	
QY 1021	GAGGGGAGCCCATGCGCCACACCTAGGGGGCAACCTGGGGGACCCAGAGCCGGCT	1080	
Db 1021	GAGGGGAGCCCATGCGCCACACCTAGGGGGCAACCTGGGGGACCCAGAGCCGGCT	1080	

Qy	1081	GTCTAGTGCCTTGGCGCGTGTGCTCGCTTTCTGAGTGGCGCTTCTCTTTGG	1140
Db	1081	GTCTAGTGCCTTGGCGCGTGTGCTCGCTTTCTGAGTGGCGCTTCTCTTTGG	1140
Qy	1141	GGCCCTGGTTACTCTTCAGCGAAATCTCCCTCATCTCTGATGTGTGAACAATCTCTCT	1200
Db	1141	GGCCCTGGTTACTCTTCAGCGAAATCTCCCTCATCTCTGATGTGTGAACAATCTCTCT	1200
Qy	1201	CCGGCACTGGGAGGCACTTCCCGCAGCCCGCTGTGGCGGCTGGCCACCTCCACCC	1260
Db	1201	CCGGCACTGGGAGGCACTTCCCGCAGCCCGCTGTGGCGGCTGGCCACCTCCACCC	1260
Qy	1261	AACCTTCAGCAGCTTGGAGCTGGAGCCAGAGCCAGCAGCCGTCGAGCCAGGGG	1320
Db	1261	AACCTTCAGCAGCTTGGAGCTGGAGCCAGAGCCAGCAGCCGTCGAGCCAGGGG	1320
Qy	1321	AGCCGACCGCATCTCATCGGCTGCTGTGGCCATCATCTGCTCTGCTGCTCATC	1380
Db	1321	AGCCGACCGCATCTCATCGGCTGCTGTGGCCATCATCTGCTCTGCTGCTCATC	1380
Qy	1381	ATTGCCCTCATGCTCTGGCGGCTGCACTGGCGCAGGCTCTCTCAGCAAGGCTGAACGGAG	1440
Db	1381	ATTGCCCTCATGCTCTGGCGGCTGCACTGGCGCAGGCTCTCTCAGCAAGGCTGAACGGAG	1440
Qy	1441	GTGTTGGAAGAGAGCTGACGGTTCACTCTCTGCTGGGACACTATCTCATCAAC	1500
Db	1441	GTGTTGGAAGAGAGCTGACGGTTCACTCTCTGCTGGGACACTATCTCATCAAC	1500
Qy	1501	AACGCGCCAGGCTTAGAGAGCCACCCCGTACAGAGAGCCCGGCTCGTGGGAATCCG	1560
Db	1501	AACGCGCCAGGCTTAGAGAGCCACCCCGTACAGAGAGCCCGGCTCGTGGGAATCCG	1560
Qy	1561	CCCGCTCGCTCTGCTGCTCCCAATGGCTTGCCTGCTCTGCTCTCAATCCAGCCTAC	1620
Db	1561	CCCGCTCGCTCTGCTGCTCCCAATGGCTTGCCTGCTCTGCTCTCAATCCAGCCTAC	1620
Qy	1621	CGCGCTCTTCTGGCCACTTACGCGCGTCCCGCTCGAGGCGCCGCGCCCCCACCGCGC	1680
Db	1621	CGCGCTCTTCTGGCCACTTACGCGCGTCCCGCTCGAGGCGCCGCGCCCCCACCGCGC	1680
Qy	1681	TGGGCAAAACCCACCAACAGCGGCTACAGTGGGACTATATGAGGCTGAGAAGCA	1740
Db	1681	TGGGCAAAACCCACCAACAGCGGCTACAGTGGGACTATATGAGGCTGAGAAGCA	1740
Qy	1741	GGCGCCCGCTCTGCCCCCAGCTTCCCGCAGACAGGCTCCCCATTATSCGAGAGCTGAC	1800
Db	1741	GGCGCCCGCTCTGCCCCCAGCTTCCCGCAGACAGGCTCCCCATTATSCGAGAGCTGAC	1800
Qy	1801	ATTGTTACCTGAGGGGCTCACGGGGGCAACACCTATGCTGTGCGCTGCACTCGCGTTCAAG	1860
Db	1801	ATTGTTACCTGAGGGGCTCACGGGGGCAACACCTATGCTGTGCGCTGCACTCGCGTTCAAG	1860
Qy	1861	GGGGCAGTGGGGATGGGCCCCCAGAGTGGATTCCTCGACTCGACTCCGCTTCAAG	1920
Db	1861	GGGGCAGTGGGGATGGGCCCCCAGAGTGGATTCCTCGACTCGACTCCGCTTCAAG	1920
Qy	1921	GAGAAGCTTGGCAGGGCCAGTTTGGGGAGGTCACCTGTGTGAGTTCGACAGCCCTCAA	1980
Db	1921	GAGAAGCTTGGCAGGGCCAGTTTGGGGAGGTCACCTGTGTGAGTTCGACAGCCCTCAA	1980
Qy	1981	GATCTGTGCTGTGATTTCCCGCTTAATGTGGGTAAAGGACACCTTTGCTGGTAGCT	2040
Db	1981	GATCTGTGCTGTGATTTCCCGCTTAATGTGGGTAAAGGACACCTTTGCTGGTAGCT	2040
Qy	2041	GTCAGATCTTACGCCCAGATGCCACCAAGATGCCAGGAATGATTTCTGAAAGAGGTG	2100
Db	2041	GTCAGATCTTACGCCCAGATGCCACCAAGATGCCAGGAATGATTTCTGAAAGAGGTG	2100
Qy	2101	AAGATCATGTGAGGCTCAAGGACCCAAACATCATTCGGCTGCTGGGCTGTGTGTCAG	2160
Db	2101	AAGATCATGTGAGGCTCAAGGACCCAAACATCATTCGGCTGCTGGGCTGTGTGTCAG	2160

Query Match									
Best Local Similarity 94.5%; Score 3438.6; DB 9; Length 3738;									
Matches 3569; Conservative 0; Mismatches 9; Indels 85; Gaps 4;									
polyA_site 3738									
BASE COUNT 698 a 1140 c 1123 g 777 t									
ORIGIN									
QY	17	GTGTGACTTTGAAGGAATGCCAAGAGATGCTGCCCCACCCCTTATGGCCCGAGGATCAG	76						
Db	77	GTGTGACTTTGAAGGAATGCCAAGAGATGCTGCCCCACCCCTTATGGCCCGAGGATCAG	136						
QY	77	GAGCTATGGACACAGAGCCCTGTCACTTTACTGTCTGTCTGTCTGTGTGGCAAGTGGAG	136						
Db	137	GAGCTATGGACACAGAGCCCTGTCACTTTACTGTCTGTCTGTGTGGCAAGTGGAG	196						
QY	137	ATGCTGACATGAAGGACATTTTGATCTTCCCAAGTCCCGCTATGCCCTGGGCATGCAGG	196						
Db	197	ATGCTGACATGAAGGACATTTTGATCTTCCCAAGTCCCGCTATGCCCTGGGCATGCAGG	256						
QY	197	ACCGGACCATCCACAGACAGTGCATCTCTGCTTCCAGCTCCCTGGTCAAGTTCACACTGCCG	256						
Db	257	ACCGGACCATCCACAGACAGTGCATCTCTGCTTCCAGCTCCCTGGTCAAGTTCACACTGCCG	316						
QY	257	CCCGCCACAGAGGTTGGAGAGAGTGCAGGGGATGGGCGCTGGTCCCGCCAGGGTCCG	316						
Db	317	CCCGCCACAGAGGTTGGAGAGAGTGCAGGGGATGGGCGCTGGTCCCGCCAGGGTCCG	376						
QY	317	TGTTTCCCAAGGAGGAGTACTTCCAGGTGGATCTACAAGCACTGCACCTGGTGGCTC	376						
Db	377	TGTTTCCCAAGGAGGAGTACTTCCAGGTGGATCTACAAGCACTGCACCTGGTGGCTC	436						
QY	377	TGTTGTCACACCCAGGACGATGCTCCGGGGCTTGGCAAGGAGTCTCCCGAGCTACC	436						
Db	437	TGTTGTCACACCCAGGACGATGCTCCGGGGCTTGGCAAGGAGTCTCCCGAGCTACC	496						
QY	437	GGCTGGCTTACTCCCGGATGCTGCGCGTGGATGGGCTTGGCAAGGACCGCTGGGCTCAGG	496						
Db	497	GGCTGGCTTACTCCCGGATGCTGCGCGTGGATGGGCTTGGCAAGGACCGCTGGGCTCAGG	556						
QY	497	AGCTGATCTCAGCAATAGGACCTTGAGGAGTGGTCTGAAGCACTTGGGCCCCCA	556						
Db	557	AGTGTATCTCAGCAATAGGACCTTGAGGAGTGGTCTGAAGCACTTGGGCCCCCA	616						
QY	557	TGCTTGCCGACTGGTTCGCTTCAACCCCGGCTGACCGGGTCAAGCGTCTGTCTGC	616						
Db	617	TGCTTGCCGACTGGTTCGCTTCAACCCCGGCTGACCGGGTCAAGCGTCTGTCTGC	676						
QY	617	GGTAGAGCTTATGGCTGCTCTGAGGAGTGGACTCTCTGTCTTACACCGCCCTGTGG	676						
Db	677	GGTAGAGCTTATGGCTGCTCTGAGGAGTGGACTCTCTGTCTTACACCGCCCTGTGG	736						
QY	677	GCACAGCAATGTATTTATCTGAGCCCGTGTACCTCAACCACTCCACTATGACGACATA	736						
Db	737	GCACAGCAATGTATTTATCTGAGCCCGTGTACCTCAACCACTCCACTATGACGACATA	796						
QY	737	CCGTGGCGGACTGCAGTATGGGGTCTGGCGCAGCTGGCAGATGGTGGTGGGCTGG	796						
Db	797	CCGTGGCGGACTGCAGTATGGGGTCTGGCGCAGCTGGCAGATGGTGGTGGGCTGG	856						
QY	797	ATGACTTTAGGAAGTCAAGAGCTCGGGTCTGGCCAGGCTATGACTATGTGGGATGA	856						
Db	857	ATGACTTTAGGAAGTCAAGAGCTCGGGTCTGGCCAGGCTATGACTATGTGGGATGA	916						
QY	857	GCAACACAGCTTCTCCAGTGGCTATGTGGAGATGGAGTTTGACCGGCTGAGGG	916						
Db	917	GCAACACAGCTTCTCCAGTGGCTATGTGGAGATGGAGTTTGACCGGCTGAGGG	976						
QY	917	CCCTCAGGCTATGCAGGCTCACTTAACAACATGCACACGCTGGGAGCCGCTGCTGCTG	976						
Db	977	CCCTCAGGCTATGCAGGCTCACTTAACAACATGCACACGCTGGGAGCCGCTGCTGCTG	1036						

QY	977	CGGGGTGGAATGTGCTTCCGGCTGGCCCTGCCATGGCTGGGAGGGAGCCCATGC	1036						
Db	1037	CGGGGTGGAATGTGCTTCCGGCTGGCCCTGCCATGGCTGGGAGGGAGCCCATGC	1096						
QY	1037	GCCACAACCTAGGGGCAACCTCTGGGGACCCAGAGCCGGGGCTGTCTCAGTCCCTCTTG	1096						
Db	1097	GCCACAACCTAGGGGCAACCTCTGGGGACCCAGAGCCGGGGCTGTCTCAGTCCCTCTTG	1156						
QY	1097	CGGGCCGTGTGGCTTCTCTCAGTGGCGCTTCTTTTGGGGGGCCCTGGTTACTCT	1156						
Db	1157	CGGGCCGTGTGGCTTCTCTCAGTGGCGCTTCTTTTGGGGGGCCCTGGTTACTCT	1216						
QY	1157	TCACGGAATCTCCTTCATCTCTGATGTGTGAACAAATTCCTCTCCGGCACTGGGAGGA	1216						
Db	1217	TCACGGAATCTCCTTCATCTCTGATGTGTGAACAAATTCCTCTCCGGCACTGGGAGGA	1276						
QY	1217	CTTTCCCGCCAGCCCTGTGTGGCGCTTGGCCCACTCCACCAACTTCAGCAGCTTG	1276						
Db	1277	CTTTCCCGCCAGCCCTGTGTGGCGCTTGGCCCACTCCACCAACTTCAGCAGCTTG	1336						
QY	1277	AGCTGGAGCCAGAGGCAAGCCGTGGCCCAAGCCGAGGGGAGCCGACCGCATCC	1336						
Db	1337	AGCTGGAGCCAGAGGCAAGCCGTGGCCCAAGCCGAGGGGAGCCGACCGCATCC	1396						
QY	1337	TCATCGGCTGCCTGGTGGCCATCATCTGCTGTCTGTCTATCATTTGCCCTCATGCTCT	1396						
Db	1397	TCATCGGCTGCCTGGTGGCCATCATCTGCTGTCTGTCTATCATTTGCCCTCATGCTCT	1456						
QY	1397	GGCGCTGCACTGGCGCAGCTCTCAGCAAGGCTGAACGAGGGTGTGGAAGAGGAGC	1456						
Db	1457	GGCGCTGCACTGGCGCAGCTCTCAGCAAGGCTGAACGAGGGTGTGGAAGAGGAGC	1516						
QY	1457	TGAGGTTTCACTCTCTGCTGGGACACTATCTCTATCAACACCGCCCAAGTCTTA	1516						
Db	1517	TGAGGTTTCACTCTCTGCTGGGACACTATCTCTATCAACACCGCCCAAGTCTTA	1576						
QY	1517	GAGAGCACCCCGTACAGAGAGCCCGGCTCGTGGGAATCCGCCCACTCGCTCCCT	1576						
Db	1577	GAGAGCACCCCGTACAGAGAGCCCGGCTCGTGGGAATCCGCCCACTCGCTCCCT	1636						
QY	1577	GTGTCCCAATGGCTGTGCTGTCTTCCCAATCCAGCTTACCGCTCTCTCTGGCCA	1636						
Db	1637	GTGTCCCAATGGCTGTGCTGTCTTCCCAATCCAGCTTACCGCTCTCTCTGGCCA	1696						
QY	1637	CTTAGCCCTTCCCGCTCGAGAGCCCGGCCCCCACACCCCGCTGGGCCAAACCCACCA	1696						
Db	1697	CTTAGCCCTTCCCGCTCGAGAGCCCGGCCCCCACACCCCGCTGGGCCAAACCCACCA	1756						
QY	1697	ACACCCAGGCTTACAGTGGGACTATATGAGGCTTGAGAGCCAGGCGCCCTTCTGC	1756						
Db	1757	ACACCCAGGCTTACAGTGGGACTATATGAGGCTTGAGAGCCAGGCGCCCTTCTGC	1816						
QY	1757	CCCCACCTCCCAAGACAGCTCCCATTTATGCCAGGCTGACATTTTACCTTCGAGG	1816						
Db	1817	CCCCACCTCCCAAGACAGCTCCCATTTATGCCAGGCTGACATTTTACCTTCGAGG	1876						
QY	1817	CGCTACCCGGGGCAACACCTATGCTGTGCTGCACTGCCCCAGGCGGAGTGGGGATG	1876						
Db	1877	CGCTACCCGGGGCAACACCTATGCTGTGCTGCACTGCCCCAGGCGGAGTGGGGATG	1936						
QY	1877	GGCCCCCAGAGTGGATTTCCCTCGATCTCGACTCCGCTTCAAGAGAACCTTGGGAGG	1936						
Db	1937	GGCCCCCAGAGTGGATTTCCCTCGATCTCGACTCCGCTTCAAGAGAACCTTGGGAGG	1996						
QY	1937	GCCAGTTTGGGAGGTCACCTGTGTGAGTGCACAGCCCTCAAGATCTGCTCAGTCTTG	1996						
Db	1997	GCCAGTTTGGGAGGTCACCTGTGTGAGTGCACAGCCCTCAAGATCTGCTTAGTCTTG	2056						
QY	1997	ATTTCCTCCCTTAATGTGCGTAAGGACACCTTTGCTGGTAGCTGTCAAGATCTTAGCGC	2056						
Db	2057	ATTTCCTCCCTTAATGTGCGTAAGGACACCTTTGCTGGTAGCTGTCAAGATCTTAGCGC	2116						
QY	2057	CAGATGCCACCAAGAAATGCCAGGAAATTTCTTGAAAGAGGTGAAGATCATGTGAGGC	2116						

|||||
Db 2117 CAGATGCCAACAAATGCCAGGAATCATTTCTCTGAAGAGGTGAAGATCATGCTCGAGGC 2176
QY 2117 TCAGAGGCCAAACATCATTTCCGCTGCTGGCGTGTGTGTCAGAGGAGGACCCCTCTGCA 2176
Db 2177 TCAAGGACCCAAACATCATTTCCGCTGCTGGCGTGTGTGTCAGAGGAGGACCCCTCTGCA 2236
QY 2177 TGATTACTGACTACATGGAGACGGCCACCTCAACCACTTCTCCTCAGTGCACCACTGG 2236
Db 2237 TGATTACTGACTACATGGAGACGGCCACCTCAACCACTTCTCCTCAGTGCACCACTGG 2296
QY 2237 AGGACAGGCGAGGCGGCGCCCTGGGACGGCGAGGCTCGGCATCGGCTATCTGCGCA 2296
Db 2297 AGGACAGGCGAGGCGGCGCCCTGGGACGGCGAGGCTCGGCATCGGCTATCTGCGCA 2356
QY 2297 GCTACCCAAATGCTGCTGATGTGCGAGCCAGATCGGCTCGGCATCGGCTATCTGCGCA 2356
Db 2357 GCTACCCAAATGCTGCTGATGTGCGAGCCAGATCGGCTCGGCATCGGCTATCTGCGCA 2416
QY 2357 CACTCAACTTTGTATCGGGACCTGGCCACGCGGAACCTGCTAGTTGGGAAATTTCA 2416
Db 2417 CACTCAACTTTGTATCGGGACCTGGCCACGCGGAACCTGCTAGTTGGGAAATTTCA 2476
QY 2417 CCATCAAAATCGCAGACTTTGGCATGAGCCGGAACCTCTATGCTGGGACTATTACCGTG 2476
Db 2477 CCATCAAAATCGCAGACTTTGGCATGAGCCGGAACCTCTATGCTGGGACTATTACCGTG 2536
QY 2477 TGCAGGCGCGGCGAGTGTGCTCCATCCGCTGGATGGCTGGGAGTGCATCTCATGGGA 2536
Db 2537 TGCAGGCGCGGCGAGTGTGCTCCATCCGCTGGATGGCTGGGAGTGCATCTCATGGGA 2596
QY 2537 AGTTCACCACTGCGAGTGTGCTGGGCTTTGGTGTGACCGCTGGGAGTGTGATGC 2596
Db 2597 AGTTCACCACTGCGAGTGTGCTGGGCTTTGGTGTGACCGCTGGGAGTGTGATGC 2656
QY 2597 TCTGTAGGCGCCAGCCCTTTGGCAGCTTCACCGAGCAGGTCTGGAATCACACATCCAG 2656
Db 2657 TCTGTAGGCGCCAGCCCTTTGGCAGCTTCACCGAGCAGGTCTGGAATCACACATCCAG 2716
QY 2657 AGTTCCTCCGGACAGGCGCGAGTGTACCTGTCCCGCGGCTGCTCCCGCGAGG 2716
Db 2717 AGTTCCTCCGGACAGGCGCGAGTGTACCTGTCCCGCGGCTGCTCCCGCGAGG 2776
QY 2717 GCCTATATGAGCTGATGCTTGGTGTGCTGGAGCGGAGTCTGAGCAGCGACCCCTTTT 2776
Db 2777 GCCTATATGAGCTGATGCTTGGTGTGCTGGAGCGGAGTCTGAGCAGCGACCCCTTTT 2836
QY 2777 CCGAGCTGCATCGGTTCTTGGCAGAGGATGACCTCAACACCGGTGTGAATCACACATCCAG 2836
Db 2837 CCGAGCTGCATCGGTTCTTGGCAGAGGATGACCTCAACACCGGTGTGAATCACACATCCAG 2896
QY 2837 CTGCGCCCTCCCTCAGGAGTGTATCAGGGGAGCCAGTGCACCTAAACACAGAGGACACA 2896
Db 2897 CTGCGCCCTCCCTCAGGAGCGCATTCAGGGGAGCCAGTGCACCTAAACACAGAGGACACA 2956
QY 2897 ATGGCAGCTCTG-CCCTTCCCTCCCGACAGCCCATCACCCTTAATAGGAGCAGTGTAGAC 2955
Db 2957 ATGGCAGCTCTGCCCCCTCCCGACAGCCCATCACCCTTAATAGGAGCAGTGTAGAC 3016
QY 2956 TGC----- 2959
Db 3017 TGCAGGCTGGGCGCCAGGAGGCTGATGCCCTTCTCCCTTCTCCTGGACACACTCTCA 3076
QY 2959 -----AGAGCCCTGTGCCCCACCCAGCTGTCTGTGG 2993
Db 3077 TGTCCCTTCTCTGTTCTTCTTCTAGAGGCCCTGTGCCCCACCCAGCTGTCTGTGG 3136
QY 2994 ATGGGATCTCTCCACCCCTCTAGCCATCCCTTGGGAGGCTGGGAGGAGGAGGAGGATATAGG 3053
Db 3137 ATGGGATCTCTCCACCCACCTCTAGCCATCCCTTGGGAGGCTGGGAGGAGGAGGAGGATATAGG 3196
QY 3054 ATAGACACTGGACATGGCCCATTTGGAGCACTTGGGCCCCCACTGGACAACTGATTCCTG 3113
|||||

Db 3197 ATAGACACTGGACATGGCCCATTTGGAGCACCTGGGCCCACTGGACAACTGATTCCTG 3256
QY 3114 GAGAGGTGGCTGGCGCCAGCTTCTCTCTGTACACACTGGACCCCACTGGGTGA 3173
Db 3257 GACAGTGGCTGGCGCCCACTTCTCTCTGTACACACTGGACCCCACTGGGTGA 3316
QY 3174 GAATCGGGGTGAGGAGGACAAAGAGGAGGAGAAATGTTTCTTGTGCTGTCTCTGT 3233
Db 3317 GAATCGGGGTGAGGAGGACAAAGAGGAGGAGAAATGTTTCTTGTGCTGTCTCTGT 3376
QY 3234 ACTTGTCTCAGCTTGGGCTTCTCTCTCTCATCACCCTGAAACACTGGACCTGGGGTA 3293
Db 3377 ACTTGTCTCAGCTTGGGCTTCTCTCTCTCATCACCCTGAAACACTGGACCTGGGGTA 3436
QY 3294 GCGCGCGCCAGCCCTCAGTCA-CCCCCACTTCCCACTGCGAGTCTTCTAGTAGAACTT 3352
Db 3437 GCGCGCGCCAGCCCTCAGTCA-CCCCCACTTCCCACTGCGAGTCTTCTAGTAGAACTT 3496
QY 3353 CTCTAAGCCTATACGTTTCTGTGGAGTAAATATTGGGATTGGGGGAAAGAGGGAGCAAC 3412
Db 3497 CTCTAAGCCTATACGTTTCTGTGGAGTAAATATTGGGATTGGGGGAAAGAGGGAGCAAC 3556
QY 3413 GCGCATAGCCCTTGGGTTGGACATCTCTAGTGTAGCTGCCACATTTTCTTATAT 3472
Db 3557 GCGCATAGCCCTTGGGTTGGACATCTCTAGTGTAGCTGCCACATTTTCTTATAT 3616
QY 3473 CACTTGGGTTTGTACATTTTGGGGGAGAGACACAGATTTTACACTATATATGAC 3532
Db 3617 CACTT-GGGTTTGTACATTTTGGGGGAGAGACACAGATTTTACACTATATATGAC 3675
QY 3533 CTAGCTTGGAGCAATTTAATCCCTGCTAGGAGTAAATAAGGTTGAGTTTC 3592
Db 3676 CTAGCTTGGAGCAATTTAATCCCTGCTAGGAGTAAATAAGGTTGAGTTTC 3735
QY 3593 CAC 3595
Db 3736 CAC 3738
|||||

RESULT 5

HUMCAK 3803 bp mRNA PRI 11-MAY-1994
LOCUS Homo sapiens tyrosine protein kinase (CAK) gene, complete cds.
DEFINITION L20817
ACCESSION L20817
VERSION 9306474
KEYWORDS L20817.1 GI:306474
SOURCE A5-antigen; factor V; factor VIII; tyrosine protein kinase.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3803)
AUTHORS Perez, J.L., Shen, X., Finkernagel, S., Sciorra, L., Jenkins, N.A.,
Gilbert, D.J., Copeland, N.G. and Wong, I.W.
TITLE Identification and chromosomal mapping of a receptor tyrosine
kinase with a putative phospholipid binding sequence in its
ectodomain
JOURNAL Oncogene 9, 211-219 (1994)
MEDLINE 9413417
REFERENCE 2 (bases 1 to 3803)
AUTHORS Johnson, J.D.
TITLE Direct Submission
JOURNAL Submitted (17-FEB-1993) J.D. Johnson
REFERENCE 3 (sites)
AUTHORS Johnson, J.D., Edman, J.C. and Rutter, W.J.
TITLE A receptor tyrosine kinase found in breast carcinoma cells has an
extracellular discoidin I-like domain
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 90, 5677-5681 (1993)
MEDLINE 93296201
REFERENCE 4 (sites)
AUTHORS Wong, I.
TITLE Direct Submission
JOURNAL Submitted (06-JUL-1993) T. Wong

Ds 1589 CTGGGACACTATCTCATCAACACCGCCGAGGTCCTTAGAGAGCCACCCCGGTACAGG 1648
Qy 1538 AGCCCGGCTCGTGGGAATCCGCCACTCGCTCCCTGTGTGTCCCAATGCTCTGGGT 1597
Ds 1649 AGCCCGGCTCGTGGGAATCCGCCACTCGCTCCCTGTGTGTCCCAATGCTCTGGGT 1708
Qy 1598 TGTGCTCTCGAATCGAGGCTACGGCTTCTTCTGGGCACATTACGCCGCTCCCTCGAG 1657
Ds 1709 TGTGCTCTCGAATCGAGGCTACGGCTTCTTCTGGGCACATTACGCCGCTCCCTCGAG 1768
Qy 1658 GCCCGGCCCCCCCCACACCGGCTGGGCCAAACCAACACCGAGGCTACAGTGGG 1717
Ds 1769 GCCCGGCCCCCCCCACACCGGCTGGGCCAAACCAACACCGAGGCTACAGTGGG 1828
Qy 1718 ACTATATGAGGCTGAGAGCGGCGGCCCGCTTCTGCCCCACCTCCCCAGAACAGCG 1777
Ds 1829 ACTATATGAGGCTGAGAGCGGCGGCCCGCTTCTGCCCCACCTCCCCAGAACAGCG 1888
Qy 1778 TCCCCCATTTATGCGAGGCTGACATTGTTACCTGAGGGGCTACCGGGGGCAACACCT 1837
Ds 1889 TCCCCCATTTATGCGAGGCTGACATTGTTACCTGAGGGGCTACCGGGGGCAACACCT 1948
Qy 1838 ATGCTGTGCTGCTGACATGCCCCAGGGGCGAGTCGGGATGGGCCCGCCAGAGTGGATTCC 1897
Ds 1949 ATGCTGTGCTGCTGACATGCCCCAGGGGCGAGTCGGGATGGGCCCGCCAGAGTGGATTCC 2008
Qy 1898 CTGATCTGAGCTCCGCTTCAAGGAGAACTTGGGAGGGCCAGTTTGGGAGGTGCACC 1957
Ds 2009 CTGATCTGAGCTCCGCTTCAAGGAGAACTTGGGAGGGCCAGTTTGGGAGGTGCACC 2068
Qy 1958 TGTGTGAGTGCAGACGCCCTCAAGATCTGGTCAGTCTTGATTCCCTTAAATGTCGTA 2017
Ds 2069 TGTGTGAGTGCAGACGCCCTCAAGATCTGGTTAGTCTTGATTCCCTTAAATGTCGTA 2128
Qy 2018 AGGAGACCTTTGCTGTGTGCTGAGTCTCAAGATCTTACGGCCAGATGCCACCAAGATGCCA 2077
Ds 2129 AGGAGACCTTTGCTGTGTGCTGAGTCTCAAGATCTTACGGCCAGATGCCACCAAGATGCCA 2188
Qy 2078 GGAATGATTCTTGAAGAGGTGAAGATCATGTCGAGGCTCAAGACCCCAACATCATTC 2137
Ds 2189 GGAATGATTCTTGAAGAGGTGAAGATCATGTCGAGGCTCAAGACCCCAACATCATTC 2248
Qy 2138 GGTGCTGGGCTGTGTGTGAGGAGGACCCCTCTGCAATGATTACTGACTACATGGAGA 2197
Ds 2249 GGTGCTGGGCTGTGTGTGAGGAGGACCCCTCTGCAATGATTACTGACTACATGGAGA 2308
Qy 2198 ACGGCGACCTCAACCAAGTTCCTCAGTGCCACACAGCTGGAGGACAGGACCGAGGGG 2257
Ds 2309 ACGGCGACCTCAACCAAGTTCCTCAGTGCCACACAGCTGGAGGACAGGACCGAGGGG 2368
Qy 2258 CCGCTGGGACGGGAGGCTGCGAGGGGCCACCATCAGCTACCAATGCTGTGCATG 2317
Ds 2369 CCGCTGGGACGGGAGGCTGCGAGGGGCCACCATCAGCTACCAATGCTGTGCATG 2428
Qy 2318 TGGAGCCAGATGCGCTCCGGCATCGGTATCTGCGCCACACTCACTTTGTACATCGGG 2377
Ds 2429 TGGAGCCAGATGCGCTCCGGCATCGGTATCTGCGCCACACTCACTTTGTACATCGGG 2488
Qy 2378 ACCTGGCCCGGGAAGTGCCTAGTTTGGGAAATTTACCATCAAAATCGCAGACTTTG 2437
Ds 2489 ACCTGGCCCGGGAAGTGCCTAGTTTGGGAAATTTACCATCAAAATCGCAGACTTTG 2548
Qy 2438 GCATAGCGGAACCTCTATGCTGGGACTATTACCGTGTGGAGGCGGGCAGTGTGCG 2497
Ds 2549 GCATAGCGGAACCTCTATGCTGGGACTATTACCGTGTGGAGGCGGGCAGTGTGCG 2608
Qy 2498 CCATCCGCTGGATGGCTGGAGTGCATCCTCATGGGAAGTTACGACTGCGAGTGGAG 2557
Ds 2609 CCATCCGCTGGATGGCTGGAGTGCATCCTCATGGGAAGTTACGACTGCGAGTGGAG 2668
Qy 2558 TGTGGGCTTTGGTGTGACCTGTGGAGGTGCTGATGCTGTAGGGCCAGCCCTTTG 2617
Ds 2669 TGTGGGCTTTGGTGTGACCTGTGGAGGTGCTGATGCTGTAGGGCCAGCCCTTTG 2728

Qy 2618 GGAGGCTACCGAGGAGGAGGTCATCGAGAACGCGGGGAGTTCCTCCGGGACCCAGGGCC 2677
Ds 2729 GGAGGCTACCGAGGAGGAGGTCATCGAGAACGCGGGGAGTTCCTCCGGGACCCAGGGCC 2788
Qy 2678 GGAGGCTACCGAGGAGGAGGTCATCGAGAACGCGGGGAGTTCCTCCGGGACCCAGGGCC 2737
Ds 2789 GGAGGCTACCGAGGAGGAGGTCATCGAGAACGCGGGGAGTTCCTCCGGGACCCAGGGCC 2848
Qy 2738 GGTGCTGGAGCCGGGAGTCTGAGCAGCAGCACCCCTTTTCCAGCTGCACTGGTTCCTGG 2797
Ds 2849 GGTGCTGGAGCCGGGAGTCTGAGCAGCAGCACCCCTTTTCCAGCTGCACTGGTTCCTGG 2908
Qy 2798 CAGAGGATGCACTCAACACCGGTGTGAATCACATCAGCTGCCCTCTCCCTCAGGAGGTG 2857
Ds 2909 CAGAGGATGCACTCAACACCGGTGTGAATCACATCAGCTGCCCTCTCCCTCAGGAGGTG 2968
Qy 2858 ATCCAGGGAGGAGGAGTGCATGACACTAAACACAGAGAGACACATGCGACCTCTGCCCTTCCC 2917
Ds 2969 ATCCAGGGAGGAGGAGTGCATGACACTAAACACAGAGAGACACATGCGACCTCTGCCCTTCCC 3028
Qy 2918 TCCGACAGCCCATCAGCTCTTAATAGAGGAGTGCAGCTGCG----- 2959
Ds 3029 TCCGACAGCCCATCAGCTCTTAATAGAGGAGTGCAGCTGCGGGTGGGCAAGGGA 3088
Qy 2959 ----- 2959
Ds 3089 GCTGATGCCCTTCTCCCTTCTGGACACACTCTCATGTCCCTTCTTCTTCTTCTTCT 3148
Qy 2959 --AGAACCCCTCTGCGCCACACAGCTGCTCTGTGATGGGATCTCTCCACCCCTCTC 3016
Ds 3149 CTAGAACCCCTCTGCGCCACACAGCTGCTCTGTGATGGGATCTCTCCACCCCTCTC 3208
Qy 3017 TAGCCATCCCTTGGGAGGAGTGGGAGAAATATAGATAGACACTGGACATGGCCATT 3076
Ds 3209 TAGCCATCCCTTGGGAGGAGTGGGAGAAATATAGATAGACACTGGACATGGCCAAAT 3268
Qy 3077 GGAGCACCTGGGCCCTTGGCAACACTGATTCCTGGAGAGTGGCTGCGCCCCAGCT 3136
Ds 3269 GGAGAACCTGGGCCCTTGGCAACACTGATTCCTGGAGAG-----GAGACCCAGCT 3322
Qy 3137 TCTCTCTCCCTGTACACACTGGACCCCACTGGTGAGAAATCTGGGGGTGAGGAGACAA 3196
Ds 3323 TCTCTCTCCCTGTACACACTGGACCCCACTGGTGAGAAATCTGGGGGTGAGGAGACAA 3382
Qy 3197 GAAGGAGAGGAAATGTTCTTGTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3256
Ds 3383 GAAGGAGAGGAAATGTTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3442
Qy 3257 TCTCTCTCCCTGTACACACTGGACCCCACTGGTGAGAAATCTGGGGGTGAGGAGACAA 3316
Ds 3443 TCTCTCTCCCTGTACACACTGGACCCCACTGGTGAGAAATCTGGGGGTGAGGAGACAA 3502
Qy 3317 CCGCACCTCCCACTGCTGAGTCTGTAGCTAGAACTCTCTAAGCTATAGCTTTCTGTGG 3376
Ds 3503 CCGCACCTCCCACTGCTGAGTCTGTAGCTAGAACTCTCTAAGCTATAGCTTTCTGTGG 3560
Qy 3377 AGTAAATATTGGGATGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3436
Ds 3561 AGTAAATATTGGGATGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3619
Qy 3437 TCTCTAGTGTAGCTGCGACATGATTTTCTATAATCAGCTTGGGTTGTGTACATTTTGG 3496
Ds 3620 TCTCTAGTGTAGCTGCGACATGATTTTCTATAATCAGCTTGGGTTGTGTACATTTTGG 3679
Qy 3497 GGGGAGAGACAGATTTTACACTAATATATGAGCTAGCTTTGAGCAATTTTAAATCC 3556
Ds 3680 GGGGAGAGACAGATTTTACACTAATATATGAGCTAGCTTTGAGCAATTTTAAAT-CC 3737
Qy 3557 CTGCACTAGGAGGATTAATAAGGTTGAGTTTCCAAAAAATAAAAAAAAAAAAAA 3616
Ds 3738 CTGCACTAGGAGGATTAATAAGGTTGAGTTTCCAAAAAATAAAAAAAAAAAAAA 3797

QY 2798 CAGAGGATGCACTCAACACGGTGTGAATACACATCCAGCTGCCCCCTCCCTCAGGGAGTG 2857
 Db 2942 CAGAGGATGCACTCAACACGGTGTGAATACACATCCAGCTGCCCCCTCCCTCAGGGAGCG 3001
 QY 2858 ATCCAGGGGAAGCAGTGCACACTAAACAAGAGGACACATGCGACCTCTGCCCTTCCCC 2917
 Db 3002 ATCCAGGGGAAGCAGTGCACACTAAACAAGAGGACACATGCGACCTCTGCCCTTCCCC 3061
 QY 2918 TCCGACAGCCCATCACCCTCTAATAGAGGCAAGTGCAGAG----- 2964
 Db 3062 TCCGACAGCCCATCACCCTCTAATAGAGGCAAGTGCAGAGTGGGCTGGGCCACCC 3121
 QY 2964 ----- 2964
 Db 3122 AGGAGCTGATGCCCCCTTCTCCCTTCTCGACACACTCTCATGTCCCTTCTCTGTTCTT 3181
 QY 2964 -----CCCTGTCGCCCCACCCAGCTGCTCTGTGGATGGATCTCTCTCCAC 3009
 Db 3182 CTTCTCTAGAAGCCCCCTGTCGCCCCACCCAGCTGCTCTGTGGATGGATCTCTCTCCAC 3241
 QY 3010 CCTCCTCTAGCATCTCTTGGGGAAGGCTGGGGAAGATATAGATAGACACTGGACATG 3069
 Db 3242 CCTCCTCTAGCATCTCTTGGGGAAGGCTGGGGAAGATATAGATAGACACTGGACATG 3301
 QY 3070 GCCATTGGACACTTGGGCCCTACTGGACAACACTGATTCCTGGAGAGTGGCTGGGCC 3129
 Db 3302 GCCATTGGACACTTGGGCCCTACTGGACAACACTGATTCCTGGAGAGTGGCTGGGCC 3361
 QY 3130 CCCAGCTTCT 3189
 Db 3362 CCCAGCTTCT 3421
 QY 3190 AGGACAAGAGGAGAGAAATGTTTCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 3249
 Db 3422 AGGACAAGAGGAGAGAAATGTTTCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 3481
 QY 3250 GGCTTCT 3309
 Db 3482 GGCTTCT 3541
 QY 3310 CAGTACCCCT 3369
 Db 3542 CAGTACCCCT 3601
 QY 3370 TCTGTGAGTAATATTTGGGATTTGGGGGAAAGAGGAGCAACGGCCCATAGCCTTTGGGG 3429
 Db 3602 TCTGTGAGTAATATTTGGGATTTGGGGGAAAGAGGAGCAACGGCCCATAGCCTTTGGGG 3661
 QY 3430 TTGGACATCTCTAGTGTAGCTGCCACATGATTTTCTATATCATCTTGGGTTGTACA 3489
 Db 3662 TTGGACATCTCTAGTGTAGCTGCCACATGATTTTCTATATCATCTTGGGTTGTACA 3721
 QY 3490 TTTTGGGGGAGAGACACAGATTTTACACTAATATATGAGCACTAGCTTTGAGGCAATTT 3549
 Db 3722 TTTTGGGGGAGAGACACAGATTTTACACTAATATATGAGCACTAGCTTTGAGGCAATTT 3781
 QY 3550 TAATCCCTCTCAGTACGAGGTAATATAAGTTGAGTTTTCACAAAAAATAAAAAA 3609
 Db 3782 TAATCCCTCTCAGTACGAGGTAATATAAGTTGAGTTTTCACAAAAAATAAAAAA 3841
 RESULT 8
 LOCUS HUMCAKA 3692 bp mRNA PRI 31-JUL-1996
 DEFINITION Homo sapiens Cak receptor kinase mRNA, complete cds.
 ACCESSION L57508
 NID g1160924
 VERSION L57508.1 GI:1160924
 KEYWORDS receptor kinase.
 SOURCE Homo sapiens lung cDNA to mRNA.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 1 (bases 1 to 3692)
 Perez, J.L., Jing, S.O. and Wong, T.W.
 Identification of two isoforms of the Cak receptor kinase that are
 coexpressed in breast tumor cell lines
 Oncogene 12 (7), 1469-1477 (1996)
 96204002
 Location/Qualifiers
 1..3692
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /tissue_type="lung"
 /note="(vector lambda gt10)"
 193..2823
 /gene="Cak"
 193..2823
 /gene="Cak"
 /codon_start=1
 /product="receptor kinase"
 /protein_id="AAB05208.1"
 /db_xref="PIR:g1160925"
 /db_xref="GI:1160925"
 /translation="MGPEALSSLLLLLLVASGDADMKHFDPAKRYALGMQDRITPD
 SDTSASSWSDSTAARSHRSLESDGDGAWCPAGSVFPKEEYLVQDLQRLHLVALVGT
 QGRHAGGLGKEFSSRYRLSRDGRMMGWKQWQEVISGNEDPEGVWLKDLGPPMV
 ARLVRFYPRADRVMSVCLRYELKWRDGLLSTAPVGTOTVLSAVYLVNDSTYDGH
 TVGGLQYGGIQLADGVVGLDPRKSOELRWPGDYVGHNSHSFSSGYSVEMEFEDR
 LRAFOAMOVHCNNHNLGRLPGVCEFRFRGFMAGEPFRHNGNGLDPRARAV
 SVPLGVARFLQCFRFLFAGPWLFLSEISFSDVNNSSPALGGTTPPAPWPPGPP
 TNFSSLEPRGQPVAKASPTAILGLVAIILLILLIALLMLRLHRRLLSKA
 ERYLEBELTHLSVFGDTILINRPGPREPPYPRPRGNPPHAPCPVNGSAYSG
 DYMEPEKGPAPLPPPPNSVPHYAEADIVTQGVGTGNTYAVPLPAGVGGPPRV
 DFRSLRFLKELGEGFGEVLDSPDLVSLDEPLNVRKHGELLVAVKLLRDA
 TKARNDFLEKVMISRLKPNILRLGVCVQDDPLCMITDYMENDGLNOLSAHOLE
 DKAEGAPGCGQAQGTIISFPMHLHVAQIASGMRYLATLNFVHRDLATRNCLVGEN
 FTIKIDFGMSRLYAGDYRVQGRAVLPFRMNAWECILMGKFTTASDYWAFGLTWE
 VLMCRAQPFQGLTDEQVIEENAGEFFRDGQVYLSRPPACPOGLYELMLRCWRESE
 QRPFPQLHRLFAEDALNTV"
 BASE COUNT 704 a 1127 c 1102 g 759 t
 ORIGIN
 Query Match 87.1%; Score 3167.6; DB 10; Length 3692;
 Best Local Similarity 94.1%; Pred. No. 0;
 Matches 3449; Conservative 0; Mismatches 14; Indels 203; Gaps 7;
 QY 38 AGAGATGCTGCCCCACCCCTTAGGCCCCGAGGATCAGAGCTATGGGACGAGGCC 97
 Db 149 AGAGATGCTGCCCCACCCCTTAGGCCCCGAGGATCAGAGCTATGGGACGAGGCC 208
 QY 98 TGTCATCTTTACTGCTGCTCTTGGTGGCAAGTGGAGATGCTGACATGAAGGGACATT 157
 Db 209 TGTCATCTTTACTGCTGCTCTTGGTGGCAAGTGGAGATGCTGACATGAAGGGACATT 268
 QY 158 TTGATCTGCAAGTGGCTATGCCCTGGGCATGCAAGCCGACCATCCACAGACAGTG 217
 Db 269 TTGATCTGCAAGTGGCTATGCCCTGGGCATGCAAGCCGACCATCCACAGACAGTG 328
 QY 218 ACATCTCTGCTTCCAGCTCTGCTGTCAGATCCACTGCCGCCGCCGACGAGTGGAGA 277
 Db 329 ACATCTCTGCTTCCAGCTCTGCTGTCAGATCCACTGCCGCCGCCGACGAGTGGAGA 388
 QY 278 GCAGTACGCGGGATGGGGCTTGGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 397
 Db 389 GCAGTACGCGGGATGGGGCTTGGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 448
 QY 338 ACTTGCAGGTGGATCTACACAGCTGCACCTGGTGGCTCTGTGGGACCCAGGGACGCC 397
 Db 449 ACTTGCAGGTGGATCTACACAGCTGCACCTGGTGGCTCTGTGGGACCCAGGGACGCC 508
 QY 398 ATGCGGGGGCTGGGCAAGGAGTTCTCCCGGAGCTACCGGCTGCGGTTACTCCCGGATG 457
 Db 509 ATGCGGGGGCTGGGCAAGGAGTTCTCCCGGAGCTACCGGCTGCGGTTACTCCCGGATG 568


```
Db 2618 GGCAGCTCACCGAGGAGAGGTCTATCAGAAACGCGGGGAGTTCTTCCGGGACGAGGCC 2677
|
Qy 2678 GGCAGGTGTACTGTCCCGCGCGCTGCCTGCCCGAGGCGCTATATGAGCTGATGCTTC 2737
|
Db 2678 GGCAGGTGTACTGTCCCGCGCGCTGCCTGCCCGAGGCGCTATATGAGCTGATGCTTC 2737
|
Qy 2738 GGTGCTGAGCGCGGAGTCTGAGCAGCGACACACCTTTTCCAGCTGCACTCGGTTCTCTGG 2797
|
Db 2738 GGTGCTGAGCGCGGAGTCTGAGCAGCGACACACCTTTTCCAGCTGCACTCGGTTCTCTGG 2797
|
Qy 2798 CAGAGGATGCACTCAACACAGTGTGAATCACACATCCAGCTGCCCTCCCTCAGGAGTG 2857
|
Db 2798 CAGAGGATGCACTCAACACAGTGTGAATCACACATCCAGCTGCCCTCCCTCAGGAGTG 2857
|
Qy 2858 ATCCAGGGAAGGAGCTGACACTAAACAGAGGACACAATGGCACCTCTGCCCTTCCCTCC 2917
|
Db 2858 ATCCAGGGAAGGAGCTGACACTAAACAGAGGACACAATGGCACCTCTGCCCTTCCCTCC 2917
|
Qy 2918 TCCGACAGCCCATCACTCTTAATAGAGGAGTGGAGTGC----- 2959
|
Db 2918 TCCGACAGCCCATCACTCTTAATAGAGGAGTGGAGTGC----- 2959
|
Qy 2959 ----- 2959
|
Db 2978 GCTGATGCCCTTCTCCCTTCTTGGACACTCTCATGTCCTTCTTCTTCTTCTTCTTCTTCT 3037
|
Qy 2959 --AGAAGCCCTGTGCGCCACCCAGCTGGTCTGTGATGGATCCCTCTCCACCTCCCTC 3016
|
Db 3038 CTAGAGCCCTGTGCGCCACCCAGCTGGTCTGTGATGGATCCCTCTCCACCTCCCTC 3097
|
Qy 3017 TAGCCATCCCTTTGGGGAAGGTGGGGAATAATAGATAGACACTGGACATGGCCCAT 3076
|
Db 3098 TAGCCATCCCTTTGGGGAAGGTGGGGAATAATAGATAGACACTGGACATGGCCCAT 3157
|
Qy 3077 GGAGCACTGGCCCTGACACACTGGACACACTGATCTCTGGAGGTGGCTGGCCCTCCAGCT 3136
|
Db 3158 GGAGCACTGGCCCTGACACACTGGACACACTGATCTCTGGAGGTGGCTGGCCCTCCAGCT 3211
|
Qy 3137 TCTCTCTCTCTGTACACACTGGACCCACTGGCTGAGATCTGGGGGTGAGGAGGACAA 3196
|
Db 3212 TCTCTCTCTGTACACACTGGACCCACTGGCTGAGATCTGGGGGTGAGGAGGACAA 3271
|
Qy 3197 GAAGGAGAGGAAATGTTTCTGTGCTGTCTGTGCTGTCTGTGCTGTCTGTGCTGTCTGT 3256
|
Db 3272 GAAGGAGAGGAAATGTTTCTGTGCTGTCTGTGCTGTCTGTGCTGTCTGTGCTGTCTGT 3331
|
Qy 3257 TCCTCTCCATCAGCTGAACACTGGACCTGGGCTAGCCCGGCTAGCCCGGCTAGCTCAGTCAC 3316
|
Db 3332 TCCTCTCCATCAGCTGAACACTGGACCTGGGCTAGCCCGGCTAGCCCGGCTAGCTCAGTCAC 3391
|
Qy 3317 CCCCACTTCCCACTGAGTCTGTAGCTAGAACTTCTTAAGCTATAGCTTTCTGTGG 3376
|
Db 3392 CCCCACTTCCCACTGAGTCTGTAGCTAGAACTTCTTAAGCTATAGCTTTCTGTGG 3449
|
Qy 3377 AGTAAATATCGGATTTGGGGAAGAGGAGGAGCAACGCCCATAGCTTGGGTTGGACA 3436
|
Db 3450 AGTAAATATCGGATTTGGGGAAGAGGAGGAGCAACGCCCATAGCTTGGGTTGGACA 3508
|
Qy 3437 TCTCTAGTGTAGCTGCCACATGATTTTCTATATCAGTCTGGGTTGTACATTTTGG 3496
|
Db 3509 TCTCTAGTGTAGCTGCCACATGATTTTCTATATCAGTCTGGGTTGTACATTTTGG 3568
|
Qy 3497 GGGGAGAGACAGATTTTACATATATATGACCTAGCTTGGAGCAATTTTAAATCCC 3556
|
Db 3569 GGGGAGAGACAGATTTTACATATATATGACCTAGCTTGGAGCAATTTTAAAT-CC 3626
|
Qy 3557 CTGCACTAGGAGGTAATAATAAGGTTGAGTTTCCACAAAAAATAAAAAAATAAAAAA 3616
|
Db 3627 CTGCACTAGGAGGTAATAATAAGGTTGAGTTTCCACAAAAAATAAAAAAATAAAAAA 3686
|
Qy 3617 AAAAAA 3622
|
Db 3687 AAAAAA 3692
|
RESULT 9
LOCUS AF026259 3659 bp mRNA ROD 27-OCT-1997
DEFINITION Mus musculus receptor-like tyrosine kinase (Nep) mRNA, complete cds.
ACCESSION AF026259
VERSION 2564939
KEYWORDS AF026259.1 GI:2564939
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 3659)
AUTHORS Zerlin,M., Julius,M.A. and Goldfarb,M.
TITLE NEP: a novel receptor-like tyrosine kinase expressed in
proliferating neuroepithelia
JOURNAL Oncogene 8 (10), 2731-2739 (1993)
MEDLINE 93390947
REFERENCE 2 (bases 1 to 3659)
AUTHORS Zerlin,M., Julius,M.A. and Goldfarb,M.
TITLE Direct Submission
JOURNAL Submitted (22-SEP-1997) Brookdale Center for Developmental and
Molecular Biology, Mount Sinai School of Medicine, 1 Gustave Levy
Place, New York, NY 10029, USA
FEATURES
Location/Qualifiers
1..3659
/organism="Mus musculus"
/strain="129"
/db_xref="taxon:10090"
/cell_type="CCE embryonic stem cells,
differentiation-induced"
1..3659
/gene="Nep"
101..2836
/gene="Nep"
/codon_start=1
/product="receptor-like tyrosine kinase"
/protein_id="AAB81866.1"
/db_xref="PIR:g2564940"
/db_xref="GI:2564940"
/translat="MGTTLSLLLLLTIGDAMKGFDPKCRYALGMDRTIP
DSISVSSSDSTAARSRLESDDGAMCPAGVPKPEEYLQVLDRLHLVAG
TQRHAGLKEFSRSLRYSRDRGRWMDKRWGQEVISGNEGCVVLKLGPPM
VARLRYPRADRVMSVCLRVLYGLRWDLISYAPVGTQWLSVWVHLNDSTYD
GYTAGGLQGLGQADGVLDLDFRQSLRVMPGYDYVWSNQSEPTGYEMEF
AISVPLGHHVRFLOCFLEAGPWLFEISFTSDVNDSDTFPPAPWPPPTN
FSSLEPRGQOPVKAEGSPTAILIGCLVAIILLIILALMLRHLHRLISKADS
RVLEELTVHLSVPGDILLNNRGPPEPPYQPRPGTTPHSPACVPPNGSALLSN
PAYELLATYARPPGPTPAWDKPTNTQACGDYMEPEKPGAPLPPPPQNSVPH
YAERDITVLQVGTGNTYAPALPPGAGDPPRPDRPRKRFKEKLGECQEVHL
CEVEDPDLVSSDFPISVHKGHPLLVAVKILRPDANKARNDLFKEVKIMSKLDPHI
IRLLGVQVDDPLCMITDYMENGLNOFLSAROLENKATOGLSGDTESDQGPITSYM
LLHGAOIASCMRYLATLNEVHRDLATNCLVGENFTIKIADFGMSRLYAGDYRVQ
GRAVLPTRWANEICLMGKTTASDVMWAFVTLWEVLMCRSQPFGLQDEQVFNAG
EFFRQGRVYLSRPPACPTLYELMLRCWSREPEQRPPFAQLHRLADDAINTV"
194..658
/gene="Nep"
/feature="Nep"
/misc_feature
1340..1411
/feature="Nep"
/misc_feature
1922..2809
/feature="Nep"
/misc_feature
724 a 1095 c 1075 g 765 t
BASE COUNT
ORIGIN
```

Query Match	65.3%	Score 2375.8	DB 12	Length 3659
Best Local Similarity	81.5%	Pred. No. 0		
Matches 2969	Conservative	0	Mismatches 582	Indels 94
Gaps				
QY	38	AGAGATGCTGCCCCACACCCCTTAGGCCCGAGGGATCAGAGCTATGGGACCAAGAGCC	97	
DB	57	AGAGATGCTGCTCCCTCTCCATAGACCAGAGGGATCCAGAGCATGGGACAGGACCC	116	
QY	98	TGTCATCTTTACTGCTGGTG---CFTCTGGTGGCAAGTGGAGATGCTGCATGAAGGGAC	154	
DB	117	TCTCATCTCTACTGCTGCTACTCTTTGGTGACAAATGGAGATGCTGACATGAAGGGAC	176	
QY	155	ATTTTGATCCTGCCAAGTGCCGCTATGCCCTGGGCATCGAGGACCGGACCATCCACAGCA	214	
DB	177	ATTTTGACCCCTGCCAAGTGCOCCTATGCCCTGGGCATCGAGGACCGCACCATCTCTGACA	236	
QY	215	GTGACATCTCTGCTTCACGCTCCTGGTCAGATTCCACTGCGCGCCGCCACAGAGTTGG	274	
DB	237	GCATATCTCTGTGTCACGCTCCTGGTCAGACTCTACCGCTGCCCGCCACAGCAGGCTGG	296	
QY	275	AGAGCAGTGACGGGATGGGGCTGGTGCOCGCGAGGGTCGTTCTTCCAAAGGAGGAGG	334	
DB	297	AAAGCAGTGATGGAGATGGGGCTTGGTCCCTGCAAGGCCCTGTGTTCCCAAGAAGAGG	356	
QY	335	AGTACTTGCAGGTGGATCTTACAACGACTGCACCTGGTGGCTCTGGTGGGACCCAGGGAC	394	
DB	357	AGTACTTGCAGGTGGACCTTCGTAGGCTACACCTGGTGGCTCTGGTGGGACCCAGGGCC	416	
QY	395	GGCATCGCGGGGCTGGCGAAGAGATTCTCCCGAGATACCGGCTGCGTTACTCCGGG	454	
DB	417	GCCATGCTGGGGTCTGGCAAAAGATTCTCCGAAGCTATCGGTTGCGTTACTCCCGAG	476	
QY	455	ATGGTCGCGCTGGATGGCTGCGAAGGACCGCTGGGCTCAGGAGTGATCTCAGGCAATG	514	
DB	477	ATGGCGCGCGCTGGATGGACTGGAGGACCGCTGGGACAGGAGTGATTCGGGGTAAGG	536	
QY	515	AGGACCCCTGAGGGAGTGTGCTGAAGGACCTTTGGGCCCCCATGGTTGCCGACTGGTTC	574	
DB	537	AGCATCTGGGGAGTGTGCTGAAGACCTTTGGGCCCCCATGGTGGCCCGCTGGTCC	596	
QY	575	GCTTCTACCCCGGGCTGACCGGGTCTAGAGGCTGTGCTGCGGGTAGAGCTCTATGGCT	634	
DB	597	GCTTCTACCCCGGGCTGACCGGGTCTAGAGTGTGCTGCTTTCGGGTGGAGCTCTATGGCT	656	
QY	635	GCTCTGAGGGATGACTCTGCTTACCGGCCCTGTGGGCGACAGCAATGTAATTTAT	694	
DB	657	GCTCTGGGGGATGACTCTGCTATACAGCCCCGTGGGCGACACCATGCAAGTTAT	716	
QY	695	CTGAGG---CCGTGTACCTCAACGACTCCACATATGACGGACATACCGTGGGCGGACTGC	751	
DB	717	CTGAGTGATGTTACATCTCAATGATTTCCACTTACGATGGATATACTGCTGGAGGCTGC	776	
QY	752	AGTATGGGGTCTGGGCCAGCTGGCAGATGTGTGGTGGGCTGGATGACTTTTAGGAAGA	811	
DB	777	AGTATGGCGTCTGGGCCAGCTGGCAGATGGCGTGGTGGGCTGGATGATTTACAGGAAA	836	
QY	812	GTGAGAGCTCGGGTCTGGCCAGGCTATGACTATGTGGGATGGAGCAACCAAGCTTCT	871	
DB	837	GCCAGAGCTCGGGGTCTGGCCAGGCTATGACTATGTGGGATGGAGCAACCAAGCTTCC	896	
QY	872	CCAGTGGCTATGTGGAGATGGAGTTGAGTTTGAACCGCTGAGGCGCTTCCAGGCTATGC	931	
DB	897	CCAGGGCTACGTGGAGATGGAGTTTGAATTTGATCGGCTGAGAACCTTCCAGACCAATGC	956	
QY	932	AGGTCCACTGTAAACATATGCACGCTGGGAGCCGCTGCTGCCTGGCGGGTGGAAATGTC	991	
DB	957	AGGTCCACTGTAAACATATGCACACTCTGGGAGCCGCTACACAGCGGGGTGGAAATGCC	1016	
QY	992	GCTTCCGGCTGGCCCTGCCATGGCTGGGAGGGGAGCCCATGGGCCACAACTTAGGG	1051	
DB	1017	GGTTTAAAGGGGTCTGCCATGGCTGGGAGGAGAGCCCTGTCCGCCATGCTCTGGGAG	1076	

Qy	1052	GCAACTGGGGAGACCCAGAGCCCGGGGTGTCTCAGTGCCCCCTTGGCGGCGGTGGGCTC	1111
Db	1077	GCAGCCTTGGAGACCCAGAGCCCGGGCACTCTCAGTGCCCTGGTGCGCAGGTGGGC	1136
Qy	1112	GCCTTCTGCAGTGCCTCTCTTTTGGGGGCGCTGGTTACTCTTCAGCGAAATCTCCT	1171
Db	1137	GCCTTCTGCAGTGCAGATTCCTTTTGCAGGTCTTGGTTACCTTTCAGTGCAGATCTCT	1196
Qy	1172	TCATCTCTGATGTGGTGAACAAATTCCTCTCGGCACTGGGAGGCACCTTCCCGCCAGCCC	1231
Db	1197	TCATCTCAGATGTGGTGAAGACTCTCT-----GACACCTTCCGACGAGCCC	1244
Qy	1232	CTGTGGTGGCGCTGGCCCACTCCCAACAATTCAGCAGCTTGGAGCTGGAGCCAGAG	1291
Db	1245	CCTGGTGGCCACTGGCCCGCTCCGACCACTTCAGCAGCTTGGAGCTGGAGCCCGGG	1304
Qy	1292	GCAGCAGCCGTGGCCAAAGCCGAGGGAGCCGACCGCCACTCTCATCGGTGCTCTGG	1351
Db	1305	GTCAACAGCCAGTGGCAAGCGGAGGGAGCCAACTGCCATCTCATTTGGCTGGCTGG	1364
Qy	1352	TGGCCATCATCTGCTCTGCTGCATCATTTGCCCTCATGCTCTGGCGGTGCACCTGGC	1411
Db	1365	TGGCCATCATCTGCTGCTCTCATCATCGCGCTGATGCTCTGGAGGCTGCACCTGGC	1424
Qy	1412	GCAGGCTCTCAGCAAGCTGAAAGGAGGGTGTGGAAAGAGGAGCTGACGGTTCACTCT	1471
Db	1425	GCGGCTGTCTCAGCAAGGCTGACAGCGCGTGTGGAGGAGGAGCTGACGGTTCACTCT	1484
Qy	1472	CTGTCCCTGGGACACTATCTTCATCAACACCGCCGAGGTCTTAGAGAGCCACCCCGT	1531
Db	1485	CTGTCCCTGGGGACACTCTTCATCAACACCGCCGAGGACCCCGAGAGCCACCCGCT	1544
Qy	1532	ACCAGGAGCCCGGCTCTGTGGAAATCGCGCCCACTCCGCTTCTGTGCCCAATGGCT	1591
Db	1545	ACCAGGAGCCCGGCTCTGGGGACTCCACCCCTTCGACACCTCGCTCCCAACGGCT	1604
Qy	1592	CTGGTGTGCTCTCCAAATCAGAGCTACCGGCTCTTCTGGCACTTAGCGCGTCCCC	1651
Db	1605	CTGGTGTGCTCTCCAAATCGGGCTACCGGCTCTTCTGGCCACTTAGCGCGTCCCC	1664
Qy	1652	CTGAGCGCGGGCGCCGCCACACCGGCTGGGCAAAACCACCAACCCAGGCTTACA	1711
Db	1665	CTGAGCGCGGGCGCCGCCACACCGGCTGGGCAAAACCACCAACCCAGGCTTACA	1724
Qy	1712	GTGGGACTATATGAGCCCTGAGAACGACGAGCGCCCGCTTCTGCCCCACCTCCCCAGA	1771
Db	1725	GTGGGACTATATGAGCCCGAGAGCGGGTGCCTCGCTTCTACCCCCACCTCCCCAGA	1784
Qy	1772	ACAGCTGCCCATTAATGCGGAGGTGACATTTTACCTTCAGGGGGGTTCACGGGGGCA	1831
Db	1785	ACAGCTGCCCATTAATGCGGAGGTGACATTTTACCTTCAGGGGGGTTCACGGGGGCA	1844
Qy	1832	ACACTATGCTGTGCTGCACCTGCCCCAGGGGAGTCGGGGATGGCCCCCAGAGTGG	1891
Db	1845	ACACTATGCTGTGCTGCACCTGCCCCAGGGGAGTGGGGATGGCCCCCAGAGTGG	1904
Qy	1892	ATTTCCCTCGACTCTCGACTCCGCTTCAAGGAGAAAGCTTGGCGAGGGCCAGTTTGGGAGG	1951
Db	1905	ATTTCCCTCGGTCAGGCTCCGCTTCAAGGAGAAAGCTTGGCGAGGGCCAGTTTGGGAGG	1964
Qy	1952	TGCACCTGTGTGAGGTGCAGACCCCTCAAGATCTGGTCAGTCTTGTATTTCCCTTAACT	2011
Db	1965	TACACCTGTGTGAAGTAGAGGACCCGCAAGATCTGGTCAGTAGTACTTCCCTATCAGT	2024
Qy	2012	TGCGTAAGGACACCCCTTGTGCTGAGTGTCAAGATCTTACGGCCAGATGCCACCAAGA	2071
Db	2025	TGCAACAAGGACACCCCTTGTGCTGAGTGTCAAGATCTTCCGGCCAGATGCCACCAAAA	2084
Qy	2072	ATGCCAGGAATGATTTCTCGAAGAGGTTCAAGATCATGTGAGGCTCAAGGACCCAAACA	2131
Db	2085	ATGCCAGGAATGATTTCTCGAAGAGGTTAAAGATCATGTACGGCTGAAGGACCCAAACA	2144
Qy	2132	TCATTCCGGCTGCTGGGCGTGTGTGTCAGGAGGACGCCCTCTGCTGATGATCTACTGACTACA	2191

sig_peptide

178..231
/gene="ptk-3"
CDS
178..2910
/gene="ptk-3"
/codon_start=1
/product="tyrosine kinase receptor"
/protein_id="AA21089.1"
/db_xref="PID:9432481"
/db_xref="GI:432481"
/translation="MGITGLSLLLLLLVTIGADMKGHFDPAKRYALQMDRTIP
DSDISVSSWSSTAAHRSRLESDDGAWCPAGVFPKEEYLOVDLRLHLVALVG
TOGRAGGLGKEFSYRLYSRDGRRWMDKRWQEVISGNEPDGVLKLDGPPM
VARLVFRPRADRVMSVCLRVELICLWRDGLLSTAPVQTMOLSENXYLNDSTYG
YTAGLQGLQGLADVVGDLDFRQSELRVWPGDYDVGNHSHFSGYVEMFEED
RURSTQVHGNHHTLGLARLPGVECFKRPAMAWEGEPVHLAGLGSGLDPRARA
ISVPLGVHRELOCFLFAGPWLFLSEISVVDNDSSTFPAPWPPGPPPTNE
SLELEPRQGVPAKAEGSPAILIGCLVAIILLILLIALLMLRHLWRLLSRAER
VLEELTVHLSVGGDTILNRPGRPPPPQEPRTPTHPSPVNGSALLSNP
ARELLATYARPPRGPTPFAWAPPTNTQACSGDMEPERKAPLPPPPQNSVPHY
AADIIVTQGVGGNTYAPVALPPGAGVGGPRVDFPRSLRFRKEKLGEGVPHLC
EYEDQDLVTSDFPISVQKHPLLVAVILKRPDLATKNARNDFLKEVIMSLKDLNII
RLGVCVDDPDLTIDYMGNDLNOFLSAHQLENKVTQGLPDRESQDQPTISYPL
LHVGAQIASGMYRLATNFRHDLATNCLNGENTIKIADFGRNRLYAGDYRVQG
WALPIRMWAECLMGKFTASDWAEGLVLEWVLMCLRSQPGQLDEQVINEAGE
FRDQGRQVLSRPACQPTLLEMLRCWSREPEORPPFSQLHRLFLADDAINTV"
232..2907

mat_peptide

/gene="ptk-3"
/product="tyrosine kinase receptor"

misc_structure

271..735
/gene="ptk-3"
/label="discoidin-like
1087..1216

misc_structure

/gene="ptk-3"
/label="cl-r-cls-like1
1222..1339
/gene="ptk-3"
/label="cl-r-cls-like2
1135 c 1078 g 789 t

BASE COUNT

ORIGIN

Query Match 63.6%; Score 2314; DB 12; Length 3743;
Best Local Similarity 80.3%; Pred. No. 0;
Matches 291; Conservative 0; Mismatches 580; Indels 149; Gaps 16;
QY 7 TCGAGTCGAGCTTGGAGCTTGAAGAAATGCCAAGAGATGCTGCCACCCCTTAGGCC 66
DB 103 TCCTCTCAGCCCTGGCTCTTCTGCTACCTTGGAGATGCTGCCCTCTCCATAGGCA 162
QY 67 GAGGGATCAGGAGCTATGGACAGAGCCCTGTATCTTTACTGCTGCTG---CTCTTG 123
DB 163 GAGGGATCAGGAGCTATGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 162
QY 124 GTGGCAAGTGGAGATCTCAGATGAGGACATTTTGTATCTCCCAAGTGGCGCTATGCC 183
DB 223 GTGACAAATGGAGATCTGACATGAGGAGACATTTTGGACCCAGCAAGTGGCGCTATGCC 282
QY 184 CTGGGATCAGGAGCCGACCATCCAGACAGTGCATCTCTGCTTCCAGCTCCCTGGTCA 243
DB 283 CTGGGATCAGGAGCCGACCATCCAGACAGGATATCTGTGTCCAGCTCTCTGGTGC 342
QY 244 GATTCCACTGCCCGCCAGCAGGAGTGGAGAGCAGTGCAGGGGATGGGCGCTGGTGC 303
DB 343 GACTCTACCGCCCGCCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 402
QY 304 CCGCAGAGTGGGTGTTTCCAGGAGGAGAGTACTTCCAGGTGATCTACAGACAGTGC 363
DB 403 CTTGAGGAGGCTGTGTCTCCCAAGAGAGAGAGTACTTCCAGGTGATCTTCTGAGGTA 462
QY 364 CACTGTGTGCTGTGTGGGACCCAGGAGGAGTGGCGGGGCTGGGCAAGAGTTC 423
DB 463 CACTGTGTGCTGTGTGGGACCCAGGAGGAGTGGCGGGTGTGGGCAAGAGTTC 522

QY 424 TCCGGAGCTACCGGCTCGCTTACTCCCGGATGGTCCCGCTGATGGCTGGAAGAC 483
DB 523 TCCCGAAGCTACCGGTTGGCTTACTCCGAGATGGCCAGCATGGATGGAGAC 582
QY 484 CGCTGGGTTCAGGAGTGATCTCAGGCAATGAGGACCTTGGAGAGTGGTGTGAAGAC 543
DB 593 AGATGGGTTCAGGAGTGATTTCCGGGTAAAGAGATCTGGGGAGTGGTGTGAAGAT 642
QY 544 CTTGGGCCCCCATGGTGGCCGACTGGTGGCTTCTACCCCGGCTGACCCGGTCATG 603
DB 643 CTGGGCCCCCATGGTGGCCGCTGGTGGCTTCTACCCCGGCTGACCCGGTTCATG 702
QY 604 AGCTGTCTCTCGGGTAGAGCTCTATGGTGGCTCTCGAGGATGAGACTCTCTCTTAC 663
DB 703 AGTGTCTCTCTCGGGTAGAGCTCTATGGTGGCTCTCGAGGAGGAGTCTTGTCTTAT 762
QY 664 ACCGCGCTCTGGGGGAGACAAATGATTTATCTAGAGCCCTGTACCTCAACAGCTCCACC 723
DB 763 ACAGCCCCCGTGGGGCAGACCATGCAATTTATCTGAGATGGTATCTCAATGATTCAC 822
QY 724 TATGAGGACATACCGTGGCGGACTGCAGTATGAGGGTCTGGGCCAGCTGGCAGATGGT 783
DB 823 TACATGATATATCTGTGGAGGCTGCAGTACGGTGTCTGGGCCAACCTGGCAGAGCGG 882
QY 784 GTGTGGGGCTGGATGACTTTAGGAAGAGTACAGAGCTGCGGGTCTGGCCAGGCTATGAC 843
DB 883 GTGTGGGGCTGGATGATTTTACGCGCAGAGCAGAGCTGCGGGTGGCCAGGCTATGAC 942
QY 844 TATGTGGATGGACACACAGCTTCTCAGTGGCTATGTGAGATGGAGTTTGAAGTTT 903
DB 943 TATGTGGATGGACACACAGCTTCTCAGTGGCTATGTGAGATGGAGTTTGAAGTTT 1002
QY 904 GACCGCTGAGGGCTTCCAGGCTATGACAGTCCAGTGTAAACACATGCACAGCTGGGA 963
DB 1003 GATCGCTGAGGCTTTCAGAGACCATGACAGTCCAGTGTAAACACATGCACAGCTGGGA 1062
QY 964 GCGGCTGCTGCTGGGGGAGTGTGCTTCCGGGCTGGGCTTCCAGTGGCTGGGAG 1023
DB 1063 GCGGCTTCCAGCGGGTGGATGCGGTTTAAAGGGTCCGGCCATGGCTGGGA 1122
QY 1024 GGGAGCCATGCGCCACACACTTAGGGGCACTTGGGGGACCCAGAGCCCGGCTGTG 1083
DB 1123 GGAGAGCTGTACACCATGCTGGGGGTAGGCTTGGAGACCCAGAGCCAGGCGCATC 1182
QY 1084 TCAGTGGCTTGGCGGCTGTGGCTGCTTCTGAGTGGCTTCTCTCTTGTGGGG 1143
DB 1183 TCGTGGCTTGGGTGTGCTGAGTGGGCGCTTCTGAGTGGCAGATTCCTCTTTCAGGA 1242
QY 1144 CCCTGGTTCCTTCAGCAAAATCTCCTTCATCTCTGATGTGGTGAACAAATCTCTCCG 1203
DB 1243 CCCTGGTTCCTTCAGTGAATCTCTTTCATCTCGGATGTGGTGAACAGTCTCTCT 1300
QY 1204 GCACTGGAGGACCTTCCCGCCAGCCCTTGGTGGCGGCTGGCCACCTCCCAAC 1263
DB 1300 -----GACACCTTCCACAGCCCTTGGTGGCGGCTGGCCCGCTCCCAAC 1350
QY 1264 TTCAGAGCTTGGAGCTGGAGCCAGAGGCGCAGAGCCCTGGCCAGCCCGAGGAGGAGC 1323
DB 1351 TTCAGAGCTTGGAGCTGGAGCCCGGGGTACAGCCCTGGCCAGGCGGAGGAGGAGC 1410
QY 1324 CCGACGCCATCTCTATCGGCTGCCTGGTGGCCATCATCTCTCTCTCTCTCTCTCTCT 1383
DB 1411 CCAACTGCCATCTCTATTTGGCTGCTGGTGGCCATCATCTCTCTCTCTCTCTCTCTCT 1470
QY 1384 GCGCTGATCTCTGGCGGCTGCACTGGCGAGGCTTCTCAGAGAGGCTGAACAGGAGGTG 1443
DB 1471 GCGCTGATCTCTGGAGGCTGCACTGGCGCGGCTGCTCAGAGGAGGAGGAGGAGG 1530
QY 1444 TTGAAGAGAGGCTGAGGTTTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1503
DB 1531 TTGAAGAGAGGCTGAGGTTTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1590
QY 1504 CGCCAGGCTCTAGAGAGCACCACCCCTTACAGGAGCCCGGCTCTGTGGGAATCCGCC 1563

Db 1591 GCGCCAGGCCCCGAGAGCCACCCCTTACAGAAACCCCGGCTCGGGGACTTCCAAT 1650
Qy 1564 CACTCCGCTCTGTGTCCTCCCAATGGCTGTGCTGTGCTCTCAATCCAGCTTACCGC 1623
Db 1651 CACTCTGGCCCTGGTCCCAATGGCTGTGCTGTGCTCTCAATCCGGCTACCGC 1710
Qy 1624 CTCTCTTGGCCACTTACGGCCCTGCCCCCTGAGGCCCCGGCCCCCACCACCCGCTGG 1683
Db 1711 CTCCTTCTGGCCACTTACGGCCCTGCCCCCTGAGGCCCCGGCCCCCACCACCCGCTGG 1770
Qy 1684 GCCAAACCCACACACCCAGCCCTACAGTGGGAGCTATATGAGCCCTGAGAACCCAGGC 1743
Db 1771 GCCAAACCCACACACCCAGCCCTGAGTGGGAGCTATATGAGCCCTGAGAACCCAGGC 1830
Qy 1744 GCGCCGCTTCTGCCCCCCTCCCAAGACAGCTGCCCCCATATGCGGAGGCTGACATT 1803
Db 1831 GCGCCGCTTCTACCCCCACCTCCCAAGACAGCTGCCCCCATATGCGGAGGCTGACATT 1890
Qy 1804 GTTACCTTSCAGGCGTACCGGGGCAACACCTATGCTGTGCTGCACTGCCCCCCAGGG 1863
Db 1891 GTCACCTTGCAGGCGTACAGGGGCAACACCTACGCTGTGCTGCACTGCCCCCCAGGG 1950
Qy 1864 CGAGTGGGGATGGCCCCCAGAGTGATTTCCCTCGATCTCGACTCCGCTTCAAGGAG 1923
Db 1951 GCGGTTGGGGATGGCCCCCAGAGTGATTTCCCTCGGTCACGGCTCCGTTTCAAGGAG 2010
Qy 1924 AAGCTTGGGAGGCGAGTTTGGGAGGTGCACCTGTGTGAGGTGCGACAGCCCTCAAGAT 1983
Db 2011 AAGCTTGGGAGGCGAGTTTGGGAGGTGCACCTGTGTGAGGTGCGACAGCCCTCAAGAT 2070
Qy 1984 CTGTGCTAGTGTATTTCCCTTAAATGTGCTAAGGACACCTTTGCTGTGAGTGTGTC 2043
Db 2071 CTGTGCTAGTGTATTTCCCTTAAATGTGCTAAGGACACCTTTGCTGTGAGTGTGTC 2130
Qy 2044 AAGATCTTACGGCCAGATGCCACCAAGATGCCAGGATGATTTCTCTGAAGAGGTGAAG 2103
Db 2131 AAGATCTTACGGCCAGATGCCACCAAGATGCCAGGATGATTTCTCTGAAGAGGTGAAG 2190
Qy 2104 ATCATGTGAGGCTCAAGGACCCAAACATCATTCGGTGTGCTGGCGTGTGTGTCAGGAC 2163
Db 2191 ATCATGTGAGGCTCAAGGACCCAAACATCATTCGGTGTGCTGGCGTGTGTGTCAGGAT 2250
Qy 2164 GACCCCTCTGCATGATTAAGTACTACATGAGAACCGGCACTCAACCAAGTTCCTCAGT 2223
Db 2251 GACCCCTCTGCATGATTAAGTACTACATGAGAACCGGCACTCAACCAAGTTCCTCAGT 2310
Qy 2224 GCGCCAGAGCTGGAGGACAGCGCCAGGGGCCCCCTGGGGACGGCGGCTGCGCAG 2283
Db 2311 GCGCCAGAGCTGGAGGACAGCGCTCACTCAGGGGCTTCTGGGGACAGAGATCTGACCAG 2370
Qy 2284 GGGCCACCATCAGCTACCCAAATGCTGTGATGTGGCAGCCAGATCGCCTCCGGCATG 2343
Db 2371 GGGCCCAACATCAGCTACCCATGCTGTATACAGTGGGGCCAGATTCCTCTGGCATG 2430
Qy 2344 CGCTATCTGGCCACTCAACTTTGTATCTCGGACCTGGCCACCGGCACTGCTAGTT 2403
Db 2431 CGATATCTGGCCACTCAACTTTGTATCTCGGACCTGGCCACCGGCACTGCTAGTT 2490
Qy 2404 GGGGAAATTTCAACATCAAAATCCACACTTTGGCATGAGCCGGAACCTTATGCTGGG 2463
Db 2491 GGGGAAATTTCAACATCAAAATCCACACTTTGGCATGAGCCGGAATCTCTACGCTGGG 2550
Qy 2464 GACTATTACCGTGTGACGGGCGGCACTGTGCTCCCATCCCTGGATGGCTGGGAGTGC 2523
Db 2551 GATTATTACCGTGTGACGGGCGGCACTGTGCTCCCATCAGTGGATGGCTGGGAGTGT 2610
Qy 2524 ATCTCATGGGAGTTTCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2583
Db 2611 ATCTCATGGGAGTTTCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2670
Qy 2584 GAGGTCCTGATGCTGTAGGCCCCAGCCCTTTGGGAGCTCACCAGGACGAGTATC 2643

Db 2671 GAGGTGCTTATGCTCTGCAAGTCAAGCCCTTGGGGCAACTCAGACGACGAGTGTATC 2730
Qy 2644 GAGAACGCGGGGAGTTCTTCCGGACACAGGCGCGCAGGTGTACCTGTCCCGCGCGCT 2703
Db 2731 GAGAACGCGGGGAGTTCTTCCAGGACACAGGCGCGCAGGTGTATCTTATCCAGGCCACC 2790
Qy 2704 GCTCCGCGCAGGCGCTATATGAGCTGATGCTTGGTCTGAGCCGGGAGTCTGAGCAG 2763
Db 2791 GCTCCGCGCAGACCTTGTATGAGCTGATGCTCCGCTTGTGAGCGCGGAGCCGAGCAG 2850
Qy 2764 CGACACACCTTTTCCAGCTGATCGTCTGCGAGAGTGCACCTCAACACGCTGGA 2823
Db 2851 CGCGCGCTTCTCCAGCTTCTCGGTCTTGGCGGATGATGCACTCAACACAGTGTAA 2910
Qy 2824 ATACACATCCAGCTGCCC-CTCCCTCAGGAGTGCATCCAGGGGAGGAGTGCACATAA 2882
Db 2911 ACTAAGGACCTGGCAGCCCTTTTTCACCTGGAGCCATCCAGGGGAAGC--TGGACATAA 2968
Qy 2883 AACAGAGACACATGAGACCTCTGCCCTTCCCCTCCCGAC----- 2925
Db 2969 AACAGAGAAAGCTATGGCACCCCGCCACCATCTCCCGGACTTCCCATTTCCCAAGAGG 3028
Qy 2925 ----AGCCCATCACCTCTAATAGAGCAGTGCAGCTGCAGAGCCCTCT----- 2971
Db 3029 CAGTGTCTCTGAGGTGGGCTGGGCCCGCCAGGAAACCGAAGACTCTGTACACAGATAC 3089
Qy 2971 -----CGCCCAACCCAGCTGCTCTGTG----- 2992
Db 3089 GTGCTCCCTTCTGTTCTTCTCCACCTGCCAAGCTCCACCCAGCTGGCCCTGTG 3148
Qy 2993 GATGGATCTCTCACCCCTCTCTAGCATCCCTTGGGGAAGGTGGGGAAGATATAG 3052
Db 3149 GTTGGATCTCTCCCTGACTTCTTCCAGCTATCCCTAGGAAAGACAGGGCGAATGCTG 3208
Qy 3053 GATAGACCTGACATGCCCATTTGGACACCTGGGCCCACTGACACACTGATTTCT 3112
Db 3209 GGCACACACTGGACAGG-----CCACTGGACAACTGTTCT 3249
Qy 3113 GGAGAGTGGCTGCGCCCCCAGCTTCTCTCTCCCTGTACACACTGGACCCCACTGGCTG 3172
Db 3250 GGGAGGTGACGTACCCCGCAGCT---TCTTCTGTACACACTGGACCGCCAG--- 3303
Qy 3173 AGAATCTGGGGTGGAGGACAAAGAGAGGAAATGTTCTTCTGCTGCTGCTCTG 3232
Db 3303 ---AATGAGGGGAGGAGGGGACAAAGAGGCAACCCCT-----CTA 3343
Qy 3233 TACTTCTCTCAGCTTGGCTTCTTCTCTCCATCA---CCTGAACACTGGACCTGGG 3289
Db 3344 CAATGCTCTCAGCCCTGGCTTCTCTGACCCCTCCGCTCTGACACACTGGACCTGGG 3403
Qy 3290 GGTAGCCCC-----GCCCGAGCCCTCAGTCAACCCCACTTCCCACTGCACTTGT 3341
Db 3404 GGTAGCCCTGCTGTGATCTCTCAGTCAAAACCCCTCTCTCTACCTACCACTGTTGC 3463
Qy 3342 AGTAGAATTC-TCTAGCCCTATAGTTCCTGTGAGTAAATATTTGGATTTGGGGGAA 3400
Db 3464 CGTAGAATTCGCAAAAGCCCTCTCGTTCCTGTGAGTAAATATTTAGAT--GGGGAAC 3521
Qy 3401 AGAGGAGCAACGGCCATAGCTTTGGGTTGGACATCTTAGTGTAGTGCACATTGA 3460
Db 3522 AGATGAGCAGTAGCCCGAGGCCGGGGTGGG--ATTCTCTGTAGTACCACATTGG 3580
Qy 3461 TTTTCTATAACTACTTGGGTTGTACA-TTTTGGGGGAGAGACACAGATTTTACA 3519
Db 3581 TTTTCTATAACTACTTGGGTTGTACATTTTGGGGGAGAGACACAGATTTTACA 3640
Qy 3520 CTAAATATGAGCTAGCTTGGGCAATTTTAAATCCCTGCACTAGGAGGTAAATAA 3579
Db 3641 CTAAATATGAGCTAGCTTGGGCAATTTTAAATCTCTGCTCTAGGAGGTAAATAA 3700
Qy 3580 AGGTTGAGTTTCCACAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3619
Db 3701 AGTTGAGTTTGTGTT 3740

RESULT 12

180847
LOCUS 180847 1197 bp DNA PAT 20-MAR-1998
DEFINITION Sequence 7 from patent US 5709858.
ACCESSION 180847
NID 93209137
VERSION 180847.1 GI:3209137
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS Godowski, P.J., Mark, M.R. and Scadden, D.T.
TITLE Antibodies specific for Rse receptor protein tyrosine kinase
JOURNAL Patent: US 5709858-A 7 20-JAN-1998;
FEATURES Location/Qualifiers
source 1..1197
BASE COUNT 199 a 351 c 407 g 240 t
ORIGIN

Query Match 32.9%; Score 1197; DB 5; Length 1197;

Best Local Similarity 100.0%; Pred. No. 1.1e-234;
Matches 1197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 136 GATGCTGACATGAAGGACATTTTCATCTGCGAAGTCCGCTATGCCCTGGGATGCAG 195
|||||
DB 1 GATGCTGACATGAAGGACATTTTCATCTGCGAAGTCCGCTATGCCCTGGGATGCAG 60
QY 196 GACCGGACCATCCAGACAGTACATCTCTGCTTCCAGCTCTCTGCTGAGATTCACATGCC 255
|||||
DB 61 GACCGGACCATCCAGACAGTACATCTCTGCTTCCAGCTCTCTGCTGAGATTCACATGCC 120
QY 256 GCCGCCACAGCAGGTTGGAGAGCAGTACAGGGGATGGGGCTGTGCCCCCGCAGGGTCG 315
|||||
DB 121 GCCGCCACAGCAGGTTGGAGAGCAGTACAGGGGATGGGGCTGTGCCCCCGCAGGGTCG 180
QY 316 GTGTTTCCCAAGGAGGAGTACTTTCAGAGTGGATCTACACGACTGCACCTGTGGCT 375
|||||
DB 181 GTGTTTCCCAAGGAGGAGTACTTTCAGAGTGGATCTACACGACTGCACCTGTGGCT 240
QY 376 CTGGTGGGACCCAGGAGCGCATCCCGGGGCTTGGCAAGAGTCTCCCGGAGGTAC 435
DB 241 CTGGTGGGACCCAGGAGCGCATCCCGGGGCTTGGCAAGAGTCTCCCGGAGGTAC 300
QY 436 CGGCTGCGTTACTCCGGGATGGTCCCGCTGGATGGGCTGGAAGGACCGCTGGGGTCAG 495
DB 301 CGGCTGCGTTACTCCCGGATGGTCCCGCTGGATGGGCTGGAAGGACCGCTGGGGTCAG 360
QY 496 GAGGTGATCTCAGGCAATGAGGACCTCAGGGAGTGGTCTGAAGGACCTTGGGCCCCCC 555
DB 361 GAGGTGATCTCAGGCAATGAGGACCTCAGGGAGTGGTCTGAAGGACCTTGGGCCCCCC 420
QY 556 ATGGTTGCCCGACTGGTTCGCTTCTACCCCGGGCTGACCGGGTCTAGAGGCTGTGCTG 615
DB 421 ATGGTTGCCCGACTGGTTCGCTTCTACCCCGGGCTGACCGGGTCTAGAGGCTGTGCTG 480
QY 615 CGGCTAGAGCTATGCTGCTCTGGAGGATGAGTCTCTGTTACACCGCCCTGTG 675
DB 481 CGGCTAGAGCTATGCTGCTCTGGAGGATGAGTCTCTGTTACACCGCCCTGTG 540
QY 676 GGGGACACAATCTATTATCTCAGGCGCTGTACCTCAACGACTCCACCTATGACGGACAT 735
DB 541 GGGGACACAATCTATTATCTCAGGCGCTGTACCTCAACGACTCCACCTATGACGGACAT 600
QY 736 ACCGTGGGCGGACTCAGTATGGGGTCTTGGGCCAGCTGGCAGATGGTGTGGTGGGGCTG 795
DB 601 ACCGTGGGCGGACTCAGTATGGGGTCTTGGGCCAGCTGGCAGATGGTGTGGTGGGGCTG 660
QY 796 GATGACTTTAGGAGAGTACAGGAGTGGGGTCTGGCCAGGCTATGACTATGTGGGATGG 855

DB 661 GATGACTTTAGGAAGAGTCAGGAGCTGCGGGTCTGGCCAGGCTATGACTATGTGGATGG 720
QY 856 AGCAACACAGCTTCTCCAGTGGCTATGTGGAGATGAGTGTGAGTTTACCGGCTGAGG 915
DB 721 AGCAACACAGCTTCTCCAGTGGCTATGTGGAGATGAGTGTGAGTTTACCGGCTGAGG 780
QY 916 GCCTTCCAGGCTATGAGGTCCACTGTAAACAACATGCACACGCTGGGAGCCGCTGTGCT 975
DB 781 GCCTTCCAGGCTATGAGGTCCACTGTAAACAACATGCACACGCTGGGAGCCGCTGTGCT 840
QY 976 GCGGGGTGGAATGTCTCTCCGGCTGGCCCTGCCATGCCCTGGAGGGGAGCCCATG 1035
DB 841 GCGGGGTGGAATGTCTCTCCGGCTGGCCCTGCCATGCCCTGGAGGGGAGCCCATG 900
QY 1036 CGCCACAACCTAGGGGCGCACTGGGGGACCCAGAGCCCGGGCTGTCTCAGTGGCCCTT 1095
DB 901 CGCCACAACCTAGGGGCGCACTGGGGGACCCAGAGCCCGGGCTGTCTCAGTGGCCCTT 960
QY 1096 GCGGGGTGGAATGTCTCTCCGGCTGGCCCTGCCATGCCCTGGAGGGGAGCCCATG 1155
DB 961 GCGGGGTGGAATGTCTCTCCGGCTGGCCCTGCCATGCCCTGGAGGGGAGCCCATG 1020
QY 1156 TTGAGGGAATCT 1215
DB 1021 TTGAGGGAATCT 1080
QY 1216 ACCTTCCCGGACCCCT 1275
DB 1081 ACCTTCCCGGACCCCT 1140
QY 1276 GAGCTGGAGCCAGAGCCAGCAGCCGCTGGCCAGCCGAGGGAGCCGAGCCGAGCCGAGCCG 1332
DB 1141 GAGCTGGAGCCAGAGCCAGCAGCCGCTGGCCAGCCGAGGGAGCCGAGCCGAGCCGAGCCG 1197

RESULT 13

AC004211/c
LOCUS AC004211 48731 bp DNA PRI 17-JUL-1998
DEFINITION Homo sapiens clone UWG:Y67C112 from 6p21, complete sequence.
ACCESSION AC004211
NID 93327405
VERSION AC004211.1 GI:3327405
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
AUTHORS Eutheria; Primates; Catarrhini; Hominidae; Homo.
JANER, M.M., GUILLAUDEUX, T., VU, Q., KUTYAVIN, T., HARTE, H. and
GERAGHTY, D.E.
TITLE Large scale sequence analysis of the human MHC class I region
JOURNAL Unpublished (1998)
REMARK Fred Hutchinson Cancer Research Center
The Clinical Research Division
1100 Fairview Ave. N., P.O. Box 19024
Seattle, WA 98109-1024
REFERENCE 2 (bases 1 to 48731)
AUTHORS Geraghty, D.E. and Olson, M.V.
TITLE Direct Submission
JOURNAL Submitted (23-FEB-1998) Human Genome Center, University of
Washington, Box 352145, Seattle, WA 98195, USA
REFERENCE 3 (bases 1 to 48731)
AUTHORS Geraghty, D.E. and Olson, M.V.
TITLE Direct Submission
JOURNAL Submitted (17-JUL-1998) Human Genome Center, University of
Washington, Box 352145, Seattle, WA 98195, USA
REMARK Contact: Daniel E. Geraghty (geraghty@fhcr.org)
COMMENT On Jul 17, 1998 this sequence version replaced gi.2905849.
Overlapping Sequences:
5': UWG:370m23.002
3': UWG:Y67C092

Sequence Quality Assessment:

This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp.

Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

Double stranded (DS) coverage: 78.6%
DS or two chemistry coverage: 98.7%
Single stranded regions: 4

Sequence Validation:

This sequence has been validated by Multiple Complete Digest Mapping. Comparison of the experimentally derived map digest fragments with sequence-predicted fragments is given below. Small fragments with a variable cutoff (approximately 400-600bp) are not mapped and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragment groups are separated by dashed lines.

BglII	HindIII	Map	Seq	Map	Seq
6380.82	6132.00	1353.33	1317.00	16545.53	16061.00
2157.94	2073.00	21029.50	20255.00	5367.59	5115.00
2205.24	2149.00	1150.03	1118.00	8273.80	7899.00
3050.67	2955.00	9349.57	8800.00	1670.66	1633.00
7318.95	7028.00	12028.04	11426.00	7024.90	6661.00
1742.45	1717.00				
3988.58	3826.00				
6401.18	6122.00				
13937.89	13391.00				

FEATURES

Source	Location/Qualifiers
variation	1. 48731
variation	/organism="Homo sapiens"
variation	/db_xref="taxon:9606"
variation	/chromosome="6"
variation	/map="6p21"
variation	/clone="CGM1:A146A3"
variation	/sub_clone="UWGC:y67c112"
variation	/cell_line="CGM1"
variation	/clone_lib="Wash U YAC library"
variation	8
variation	/note="high quality variation versus 5' overlapping clone"
variation	832
variation	/note="high quality variation versus 5' overlapping clone"
variation	1134
variation	/note="high quality variation versus 5' overlapping clone"
variation	1220
variation	/note="high quality variation versus 5' overlapping clone"
variation	1497
variation	/note="high quality variation versus 5' overlapping clone"
variation	1648
variation	/note="high quality variation versus 5' overlapping clone"
variation	2443
variation	/note="high quality variation versus 5' overlapping clone"
variation	2583
variation	/note="high quality variation versus 5' overlapping clone"
variation	3534
variation	/note="high quality variation versus 5' overlapping clone"
variation	3542

variation
/note="high quality variation versus 5' overlapping clone"
3715
/note="high quality variation versus 5' overlapping clone"
3923
/note="high quality variation versus 5' overlapping clone"
3925
/note="high quality variation versus 5' overlapping clone"
4447. 4798
/rpt_family="Alu"
5432. 5703
/rpt_family="Alu"
/rpt_family="Alu"
complement(6331. 6504)
/rpt_family="MIR"
complement(11000. 11096)
/rpt_family="MIR"
13159. 13460
/rpt_family="Alu"
15959. 16254
/rpt_family="Alu"
16800. 16713
/rpt_family="Alu"
complement(16741. 17029)
/rpt_family="Alu"
17499. 17755
/rpt_family="Alu"
17946. 18420
/rpt_family="Alu"
18446. 18566
/rpt_family="Alu"
18620. 19230
/rpt_family="Alu"
19327. 19629
/rpt_family="Alu"
complement(19714. 19876)
/rpt_family="Alu"
19991. 20271
/rpt_family="Alu"
21408. 21698
/rpt_family="Alu"
complement(22064. 22343)
/rpt_family="Alu"
complement(22946. 23220)
/rpt_family="Alu"
complement(29480. 29776)
/rpt_family="Alu"
34016. 34143
/rpt_family="Alu"
35865. 35989
/rpt_family="Alu"
43351. 43620
/rpt_family="Alu"
44069. 44316
/rpt_family="Alu"
44359. 44637
/rpt_family="Alu"
45421. 45550
/rpt_family="Alu"
complement(46394. 46937)
/rpt_family="Alu"
complement(47953. 48515)
/rpt_family="LTR9"
48682. 48731
/rpt_family="Alu"
BASE COUNT 12618 a 13116 c 12767 g 10230 t
ORIGIN

Query Match 23.08; Score 837.2; DB 11; Length 48731;
Best Local Similarity 83.7%; Pred.No. 3.1e-161;
Matches 1063; Conservative 0; Mismatches 3; Indels 204; Gaps 2;

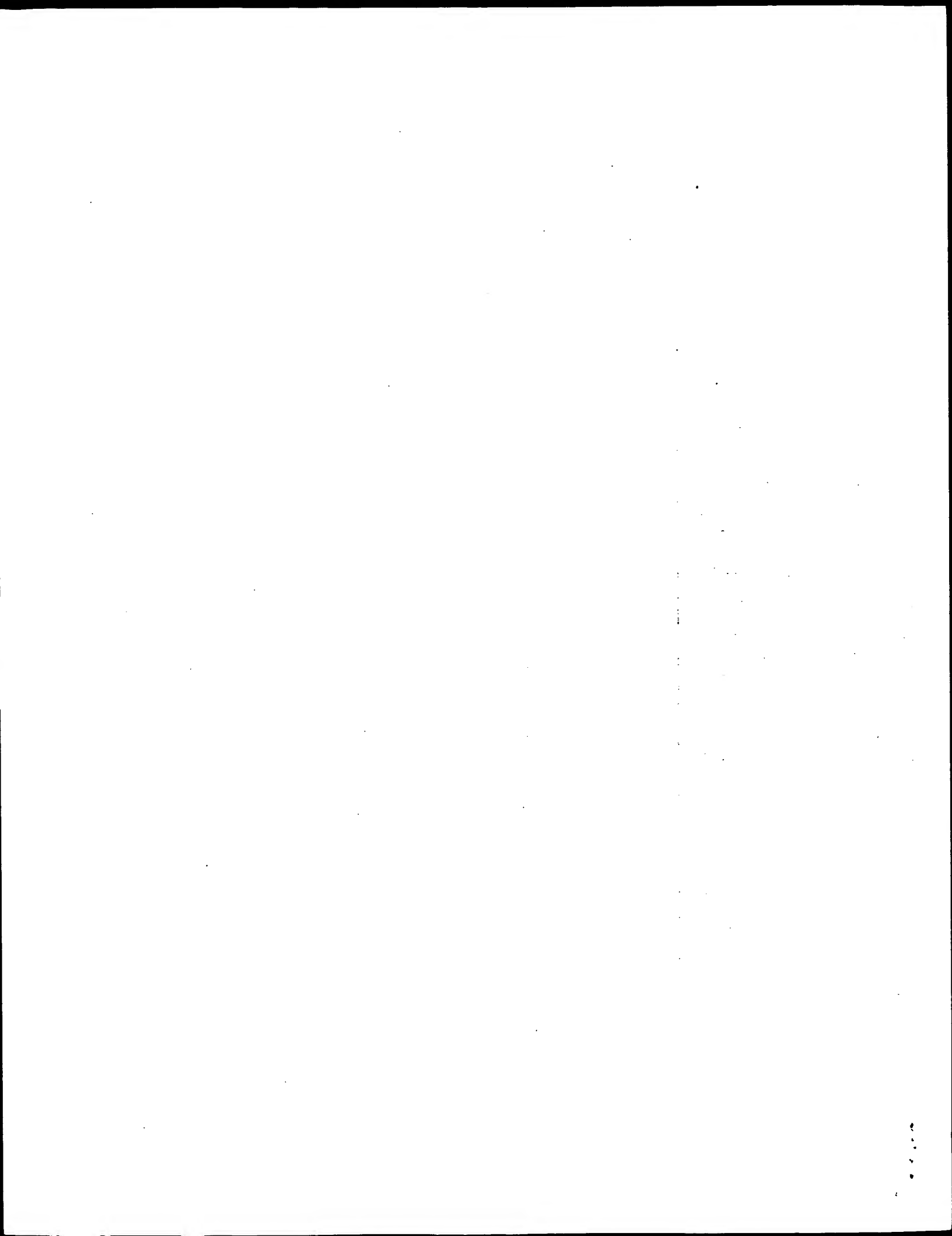
Oy 2532 GGGGAAGTTCACGACTCGAGCTGACGTGTGGCCCTTGGTGACCTGTGGAGGTCT 2591
|||||

Ddb	294	CCGACAGCCCATCACCTCTAATAGAGGCAGTGGACTCGAGGTGGGCTGGGCCCAACCCA	353
Qy	2964	-----	2964
Ddb	354	GGAGCTGATGCCCTTCCTCCCCTCCTGGACACACTCTCATGTGCCCTTCCTGTTCTTC	413
Qy	2964	-----CCCCGTGCGCCACCAGCTGGTCTCTGTGGATGGGATGCTCTCCACC	3010
Ddb	414	CTTGCTAGAAGCCCCCTGTGCGCCACCAGCTGGTCTCTGTGGATGGGATCTCTCCACC	473
Qy	3011	CTCCTCTAGCCATCCCTTTGGGAAGGTTGGGAGAAATATAGATAGACACTGCACATGG	3070
Ddb	474	CCTCTTAGCCATCCCTTTGGGAAGGTTGGGAGAAATATAGATAGACACTGCACATGG	533
Qy	3071	CCATTGGAGACCTTGGGCCCCACTGGACAACACTGATTCCTGGAGAGGTGGCTGCGCC	3130
Ddb	534	CCATTGGAGACCTTGGGCCCCACTGGACAACACTGATTCCTGGAGAGGTGGCTGCGCC	593
Qy	3131	CCAGCTTCTCTCCCTGTACACACTGGACCCCACTGGCTGAGAACTCTGGSGCTGAGGA	3190
Ddb	594	CCAGCTTCTCTCCCTGTACACACTGGACCCCACTGGCTGAGAACTCTGGSGGTGAGGA	653
Qy	3191	GSACAAGAAGGAGAGAAAATGTTCTCTGTGCTCTCTCTACTTCTCTAGCTTGG	3250
Ddb	654	GGACAAGAAGGAGAGAAAATGTTCTCTGTGCTCTCTCTACTTCTCTAGCTTGG	713
Qy	3251	GCTTCTCTCTCCATCAGCTGAACACTGGACCTGGGGGTAGCCCGCCCAAGCCCTC	3310
Ddb	714	GCTTCTCTCTCCATCAGCTGAACACTGGACCTGGGGGTAGCCCGCCCAAGCCCTC	773
Qy	3311	AGTCAACCCCACTTCCCACTGCAGTCTTGTAGCTAGAACCTTCTCTAAGCCTATACGTTT	3370
Ddb	774	AGTCAACCCCACTTCCCACTGCAGTCTTGTAGCTAGAACCTTCTCTAAGCCTATACGTTT	833
Qy	3371	CTGTGGAGTAATANTTGGAGTTGGGGGAAGAGGAGCAACGGCCCATAGCCTTGGGGT	3430
Ddb	834	CTGTGGAGTAATANTTGGAGTTGGGGGAAGAGGAGCAACGGCCCATAGCCTTGGGGT	893
Qy	3431	TGGACATCTCTAGTGTAGCTGCACATGATTTTTCTATAACTCTCTGGGTTTGTACAT	3490
Ddb	894	TGGACATCTCTAGTGTAGCTGCACATGATTTTTCTATAACTCTCTGGGTTTGTACAT	953
Qy	3491	TTTTGGGGGAGACACAGATTTTTACACTAATATATGGACCTAGCTTGGAGCAATTTT	3550
Ddb	954	TTTTGGGGGAGACACAGATTTTTACACTAATATATGGACCTAGCTTGGAGCAATTTT	1013
Qy	3551	RATCCCTGCAGTACGGCAGCTAATAAAGTTGAGTTTTCCACAAAAAARAAAAA	3609
Ddb	1014	AATCCCTGCAGTACGGCAGCTAATAAAGTTGAGTTTTCCACAAAAAARAAAAA	1072
<hr/>			
RESULT	15		
LOCUS	HSU48705	12010 bp	DNA
DEFINITION	Human receptor tyrosine kinase DDR gene, complete cds.	PRI	28-DEC-1996
ACCESSION	U48705		
NID	g1753221		
VERSION	U48705.1	GI:	1753221
KEYWORDS	.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 12010) Sakuma,S., Saya,H., Tada,M., Nakao,M., Fujiwara,T., Roth,J.A., Sawamura,Y., Shinohe,Y. and Abe,H. Receptor protein tyrosine kinase DDR is up-regulated by p53 protein FEBS Lett. 398 (2-3), 165-169 (1996)		
JOURNAL	2 (bases 1 to 12010)		
MEDLINE	97131588		
REFERENCES	2 (bases 1 to 12010) Sakuma,S., Tada,M., Saya,H., Sawamura,Y., Shinohe,Y. and Abe,H. Direct Submission		
TITLE			

QY 2532 GGGAACTTACGACTCCGAGTGCAGTGTGGGCTTTGGTGTGACCTGTGGAGTGTCT 2591
|||||
Db 10746 GGGGAAGTTACGACTCGAGTGCAGTGTGGGCTTTGGTGTGACCTGTGGAGTGTCT 10805
QY 2592 GATGCTCTGTAGGGCCAGCCCTTTGGGAGCTCACCGAGGAGCAGGTGATCGAAGCC 2651
|||||
Db 10806 GATGCTCTGTAGGGCCAGCCCTTTGGGAGCTCACCGAGGAGCAGGTGATCGAAGCC 10865
QY 2652 GGGGAGTCTTCCGGACACAGGCCCGG----- 2680
|||||
Db 10866 GGGGAGTCTTCCGGACACAGGCCCGGAGGTCAGAGTGGAGGAGGGAAGTGGTCT 10925
QY 2680 ----- 2680
Db 10926 CGAGNCGGGGAGAGAGGGGAGAGTTCATCTTGGAGACTAAAGAAATAATGTTCC 10985
QY 2680 -----CAGTGTACCTGTCCCGGCCGCTGCTGCCCGC 2713
|||||
Db 10986 TGACTCTCAACACACTTCCNCAATNCAGGTGTACCTGTCCCGGCCGCTGCTGCCCGC 11045
QY 2714 AGGSCCTATATGAGCTGATCTTCGGTGTGGAGCGGGAGTCTGAGCAGCGACACACCT 2773
|||||
Db 11046 AGGSCCTATATGAGCTGATCTTCGGTGTGGAGCGGGAGTCTGAGCAGCGACACACCT 11105
QY 2774 TTCCAGCTGCATCGGTTCTCGCAGAGATGCATCAACAGGTGTGAATCACACATC 2833
|||||
Db 11106 TTCCAGCTGCATCGGTTCTCGCAGAGATGCATCAACAGGTGTGAATCACACATC 11165
QY 2834 CAGTGCCTCCCTCAGGAGTGTCCAGGGAGGATCCAGGGAGCCAGTGACACTAAACAAAGAGGAC 2893
|||||
Db 11166 CAGTGCCTCCCTCAGGAGGATCCAGGGAGGATCCAGGGAGCCAGTGACACTAAACAAAGAGGAC 11225
QY 2894 ACAATGSCACTCTG-CCTTCCCTCCGACAGCCCATCACCTCTAATAGAGGAGTGA 2952
|||||
Db 11226 ACAATGSCACTCTGCCCCCTTCCCTCCGACAGCCCATCACCTCTAATAGAGGAGTGA 11285
QY 2953 GACTGC----- 2959
|||||
Db 11286 GACTGCAGGCTGGGCCACCCAGGAGTGTATGCCCTTCTCCCTTCTCTGGACACATC 11345
QY 2959 -----AGAGCCCTGTGCCACCCAGCTGTCTCTG 2990
|||||
Db 11346 TCATGTCCCTTCTCTTCTTCTAGAGCCCTGTGCCACCCAGCTGTCTCTG 11405
QY 2991 TGGATGGATCTCTCCACCCCTCTCTAGCCATCCCTTGGGAAGGTTGGGAGAAATAT 3050
|||||
Db 11406 TGGATGGATCTCTCCACCCCTCTAGCCATCCCTTGGGAAGGTTGGGAGAAATAT 11465
QY 3051 AGGATAGACACTGGACATGGCCCATTTGGAGCACCTGGGCCCCACTGGACAACTGATTC 3110
|||||
Db 11466 AGGATAGACACTGGACATGGCCCATTTGGAGCACCTGGGCCCCACTGGACAACTGATTC 11525
QY 3111 CTGGAGAGTGGCTGGGCCCCAGCTTCTCTCCCTGTGCACACTGGACCCCACTGGC 3170
|||||
Db 11526 CTGGAGAGTGGCTGGGCCCCAGCTTCTCTCCCTGTGCACACTGGACCCCACTGGC 11585
QY 3171 TGAGAACTCTGGGGTGGAGGAGCAGAGAGGAGAAATGTTCCCTGTGCTGCTCC 3230
|||||
Db 11586 TGAGAACTCTGGGGTGGAGGAGCAGAGAGGAGAAATGTTCCCTGTGCTGCTCC 11645
QY 3231 TGTACTTGTCTCAGCTTGGGCTTCTTCTCTCCATCACCTGAAACACTGGACCTGGGG 3290
|||||
Db 11646 TGTACTTGTCTCAGCTTGGGCTTCTTCTCTCCATCACCTGAAACACTGGACCTGGGG 11705
QY 3291 GTAGCCCCCCCCAGCCCTCAGTCA-CCCCCACTTCCCACTGCACTTGTAGTAGAA 3349
|||||
Db 11706 GTAGCCCCCCCCAGCCCTCAGTCACTCCCCCACTTCCCACTGCACTTGTAGTAGAA 11765
QY 3350 CTTCTCTAAGCCTATACGTTTCTGTGGAGTAAATTTGGGATTTGGGGAGAGAGGAGC 3409
|||||
Db 11766 CTTCTCTAAGCCTATACGTTTCTGTGGAGTAAATTTGGGATTTGGGGAGAGAGGAGC 11825

QY 3410 AACGCCCATAGCCTTGGGTTGGACATCTCTAGCTAGCTGCCACATTTGATTTTCTAT 3469
|||||
Db 11826 AACGCCCATAGCCTTGGGTTGGACATCTCTAGCTAGCTGCCACATTTGATTTTCTAT 11885
QY 3470 AATCAGCTTGGGTTGTACATTTTGGGGGAGAGACACAGATTTTACACTAATATG 3529
|||||
Db 11886 AATCAGCTT-GGGTTGTACATTTTGGGGGAGAGACACAGATTTTACACTAATATG 11944
QY 3530 GACCTAGCTTGCAGCAATTTTAAATCCCTGCACCTAGCAGGTAATAATAAGGTTGAGTT 3589
|||||
Db 11945 GACCTAGCTTGCAGCAATTTTAAATCCCTGCACCTAGCAGGTAATAATAAGGTTGAGTT 12004
QY 3590 TTCCAC 3595
|||||
Db 12005 TTCCAC 12010

Search completed: November 4, 1999, 07:31:19
Job time: 6478 sec



GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 1999, 23:46:05 ; Search time 19.79 Seconds
(without alignments)
1848.398 Million cell updates

Title: US-08-170-558-4
 Perfect score: 4903
 Sequence: 1 MGPEALSSLLLLLLVASGDA.....QRPPFSQLHRFLAEDALNTV 913

Scoring table: BLOSUM62
Searched: 122810 seqs, 40065486 residues

```
Database : PIR_60:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query			ID	Description
	Score	Match	Length		
1	4889	99.7	913	2	A48280
2	4664.5	95.1	876	2	A49508
3	4584.5	93.5	910	2	A5137
4	2415	49.3	855	2	A32621
5	2332	48.0	819	2	I48859
6	1154	23.5	220	2	I80329
7	952	19.4	248	2	A44547
8	924.5	18.9	183	2	I59442
9	661	13.5	818	1	S44098
10	659.5	13.5	821	1	S06943
11	659.5	13.5	821	1	A39667
12	658	13.4	822	1	A58853
13	653.5	13.3	799	1	TVRITB
14	645.5	13.3	825	2	A53178
15	642.5	13.1	825	1	A40026
16	639	13.0	790	1	TVHUTT
17	633.5	12.9	839	3	I73632
18	628	12.8	503	4	S23741
19	623.5	12.7	803	2	S35695
20	621	12.7	282	2	I51236
21	614	12.5	832	2	I51259
22	608.5	12.4	946	2	A47299
23	603	12.3	871	2	I48696
24	601.5	12.3	881	2	I48697
25	601.5	12.3	605	2	S67815
26	600.5	12.2	863	2	S60738
27	599	12.2	873	2	S60740
28	591.5	12.1	685	2	A48289
29	580	11.8	981	1	FOVWGM
30	546	11.1	937	2	A45082
31	544.5	11.1	402	1	TVFVUR
32	534.5	10.9	1382	1	INHUR
33	534.5	10.9	2338	2	I73957
34	534.5	10.9	2317	2	I56752
35	534	10.9	2347	1	TVRURS
36	533.5	10.9	2311	1	TVCHSR
37	530.5	10.8	1372	2	A34157
38	530	10.8	2340	2	I48310
39	528	10.8	1520	1	TVFFFA

ALIGNMENTS

RESULT 1

A48280

receptor tyrosine kinase - human

C:Species: Homo sapiens (man)

C:Date: 31-May-1996 #sequence_revision 31-May-1996 #text_change 16-Dec-1998

C:Accession: A48280

R:Johnson, J.D.; Edman, J.C.; Rutter, W.J.

Proc. Natl. Acad. Sci. U.S.A. 90, 5677-5681, 1993

A>Title: A receptor tyrosine kinase found in breast carcinoma cells has an extracellular

A:Reference number: A48280; MUID:93296201

A:Accession: A48280

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-913 <RES>

A:Cross-references: GB:L11315; NID:9403386; PID:9403387

C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; discoidin I amino-

F:30-185/Domain: discoidin I amino-terminal homology <DNI>

F:608-912/Domain: protein kinase homology <KIN>

F:616-624/Region: protein kinase ATP-binding motif

Query Match	99.7%	Score	4889	DB 2	Length	913
Best Local Similarity	99.7%	Pred. No.	1.5e-263			
Matches	910	Conservative	2	Mismatches	0	Gaps
QY	1	MGPEALSLLLLLLVASGDADMKGHPDPACRYALGMDRTIPDSDIASSSWSDSTAAR	60			
DB	1	MGPEALSLLLLLLVASGDADMKGHPDPACRYALGMDRTIPDSDIASSSWSDSTAAR	60			
QY	61	HSRLESSDGDGAWCPAGSVFPKEEYLQVDLQRLHLVALVGTQGRHAGGLGKEFSRYRL	120			
DB	61	HSRLESSDGDGAWCPAGSVFPKEEYLQVDLQRLHLVALVGTQGRHAGGLGKEFSRYRL	120			
QY	121	RYSRDGRRWGKDRWQOEYISGNDEPGVYLDKLGPPMVARLYRFYPRADRVMSVCLRV	180			
DB	121	RYSRDGRRWGKDRWQOEYISGNDEPGVYLDKLGPPMVARLYRFYPRADRVMSVCLRV	180			
QY	181	ELYGCLWRDGLLSTYAPVGTMTYLSEAYLNDSTYDCHTGVGGLOQYGGLOGLADGVGLDD	240			
DB	181	ELYGCLWRDGLLSTYAPVGTMTYLSEAYLNDSTYDCHTGVGGLOQYGGLOGLADGVGLDD	240			
QY	241	FRKSQELRVWPGDYVQWSNHSFSSGVMEFEFDRLRAFOAMOVHCNNMHTLGARLPGG	300			
DB	241	FRKSQELRVWPGDYVQWSNHSFSSGVMEFEFDRLRAFOAMOVHCNNMHTLGARLPGG	300			
QY	301	VECFRFRGPAMWEGEPMRNLGNLGDPRARAVSVPLGGRVARFLOCRLFAGPWLIFS	360			
DB	301	VECFRFRGPAMWEGEPMRNLGNLGDPRARAVSVPLGGRVARFLOCRLFAGPWLIFS	360			
QY	361	EISFISDVVNSSPALGGTFPPAPWMPGGPPPTNFSSLELEPRGQQPVAKPEGSPITALI	420			
DB	361	EISFISDVVNSSPALGGTFPPAPWMPGGPPPTNFSSLELEPRGQQPVAKPEGSPITALI	420			
QY	421	GCLVAITILLLLIITALLMLRHLWRRLLSKAERRVLEELTVHLSVPGDTILINNRPGPRE	480			
DB	421	GCLVAITILLLLIITALLMLRHLWRRLLSKAERRVLEELTVHLSVPGDTILINNRPGPRE	480			
QY	481	PPPYQEPDRPGNPHSPACVPVNGSALLLSNPAYRLLIATYARPRGCGPPTPAWAKPNT	540			
DB	481	PPPYQEPDRPGNPHSPACVPVNGSALLLSNPAYRLLIATYARPRGCGPPTPAWAKPNT	540			
QY	541	QAYSQGYMEKPGCAPILLPPPPQNSVPHYAEADITVTLQGVTTGGNTYAVPALPPGAVGDGP	600			

QY 1 MGPEALSS-LLLLLLVASGDADMKGHFDPKACRYALGMDRTIPDSISASSWSDDSTAA 59
 Db 1 MGTGTTSSLLLLLLLLLTIIGDADMKGHFDPKACRYALGMDRTIPDSISASSWSDDSTAA 60
 QY 60 RHRLESSDGDGAWCPAGSVPKKEEYLOVDLQRLHLVALVGTQGRHAGGLGKEFSRYSR 119
 Db 61 RHRLESSDGDGAWCPAGSVPKKEEYLOVDLQRLHLVALVGTQGRHAGGLGKEFSRYSR 120
 QY 120 LRYSDRGRRWGWKDRWGQEVISGNEDEPGVVLKDLGPPMVARLYRFPYPRADRVMSVCLR 179
 Db 121 LRYSDRGRRWGWKDRWGQEVISGNEDEPGVVLKDLGPPMVARLYRFPYPRADRVMSVCLR 180
 QY 180 VELYGCLWRDGLLSYAPVGTQMYLSEAYVINDSYDHTVCGLOYGLGLQADGVGLD 239
 Db 181 VELYGCLWRDGLLSYAPVGTQMYLSEAYVINDSYDHTVCGLOYGLGLQADGVGLD 240
 QY 240 DFKSKQELRWPCYDYGWNSHFSFGYVEMEFEDRLRAFAQMVHCHNNHHTLGLARLP 299
 Db 241 DFKSKQELRWPCYDYGWNSHFSFGYVEMEFEDRLRAFAQMVHCHNNHHTLGLARLP 300
 QY 300 GVECRFRGPANAWGEPVHNLGNLGDPRARAVSVPLGGVRARFLQCRFLFAGPWLFF 359
 Db 301 GVECRFRGPANAWGEPVHNLGNLGDPRARAVSVPLGGVRARFLQCRFLFAGPWLFF 360
 QY 360 SEISFISDVVNNSSPALGTFPPAPWPPGPPPTNFSSLELEPRCQOPVAKPEGSPAIL 419
 Db 361 SEISFISDVVNNSSPALGTFPPAPWPPGPPPTNFSSLELEPRCQOPVAKPEGSPAIL 416
 QY 420 IGLVAIILLLLLLIALMLRHLWRLLSKAERRVLEELTVHLSVPGDITILNRRGPR 479
 Db 417 IGLVAIILLLLLLIALMLRHLWRLLSKAERRVLEELTVHLSVPGDITILNRRGPR 476
 QY 480 EPPVQEPFRGNPNPHSAPCVNGSALLSNPAYRLLATYARPPGPPPTPAWAKPTN 539
 Db 477 EPPVQEPFRGNPNPHSAPCVNGSALLSNPAYRLLATYARPPGPPPTPAWAKPTN 536
 QY 540 TQAYSGDYMEPEKPGAPILLPPPPQNSVPHYADIVTLQVGTGNTYAVPALPGAVGDG 599
 Db 537 TQAYSGDYMEPEKPGAPILLPPPPQNSVPHYADIVTLQVGTGNTYAVPALPGAVGDG 596
 QY 600 PPRVDFPRSLRFLKELGEGQGEVHLCVDSQDLVSLDFPLNVRKGLHLLVAVKILRP 659
 Db 597 PPRVDFPRSLRFLKELGEGQGEVHLCVDSQDLVSLDFPLNVRKGLHLLVAVKILRP 656
 QY 660 DATKARNDFLKEVKIMSLKDPNIIIRLLGVCVQDDPLCMITDYMENGLDNOFTSAHOLE 719
 Db 657 DATKARNDFLKEVKIMSLKDPNIIIRLLGVCVQDDPLCMITDYMENGLDNOFTSAHOLE 716
 QY 720 DKAAEGAPGDGQAAGPTISYPMILHVAQAQASGMRYLATLNFVHRDLATNCLVGENFT 779
 Db 717 NKVTOGLPDRSDGPTISYPMILHVAQAQASGMRYLATLNFVHRDLATNCLVGENFT 776
 QY 780 IKIADFGMSRLNYAGDYRVQGRVAVLPIRWMAWECILMGKFTTASDVWAFGVTILWEVLM 839
 Db 777 IKIADFGMSRLNYAGDYRVQGRVAVLPIRWMAWECILMGKFTTASDVWAFGVTILWEVLM 836
 QY 840 CRAQPFQGLTDEQVTEENAGEFRDQGRVYLSRPPACQGLYELMLRCWSRESQRPFFS 899
 Db 837 CRAQPFQGLTDEQVTEENAGEFRDQGRVYLSRPPACQGLYELMLRCWSRESQRPFFS 896
 QY 900 QLHRELAEDALNTV 913
 Db 897 QLHRELAEDALNTV 910
 RESULT 4
 S42621
 protein-tyrosine kinase (EC 2.7.1.112) - human
 C:Species: Homo sapiens (man)
 C:Date: 07-Sep-1994 #sequence_revision 10-Nov-1995 #text_change 17-Mar-1999
 R:karn, T.; Holtrich, U.; Braeuning, A.; Boehme, B.; Wolf, G.; Ruebsamen-Waigmann, H.;
 Oncogene 8, 3433-3440, 1993

A:Title: Structure, expression and chromosomal mapping of TKT from man and mouse: a n
 A:Reference number: S42621; MUID:94067796
 A:Accession: S42621
 A:Molecule type: mRNA
 A:Residues: 1-855 <KAR>
 A:Cross-references: EMBL:X74764; NID:g433337; PID:g433338
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; discolidin I amino-
 C:Keywords: phosphotransferase; tyrosine-specific protein kinase
 F:29-185/Domain: discolidin I amino-terminal homology <DNI>
 F:561-855/Domain: protein kinase homology <KIN>
 F:569-577/Region: protein kinase ATP-binding motif

Query Match 49.3%; Score 2415; DB 2; Length 855;
 Best Local Similarity 52.1%; Pred. No. 1.2e-126;
 Matches 482; Conservative 118; Mismatches 227; Indels 98; Gaps 15;

QY 3 PEALSSLLLLLVASGDADMKGHFDPKACRYALGMDRTIPDSISASSWSDDSTAAHRS 62
 Db 5 PRMLLVLLPILS---SAKAQVNPACRYPLGMSGGQIPDEDITASSOWSESTAAYK 61
 QY 63 RLESSDGDGAWCPAGSVPKPE-EYLOVDLQRLHLVALVGTQGRHAGGLGKEFSRYSR 121
 Db 62 RLESSDGDGAWCPPEIPVPPDLEFLQIDLHTLFTLVGTQGRHAGGLGKEFSRYSR 121
 QY 122 YSRDGRWRWGWKDRWGQEVISGNEDEPGVVLKDLGPPMVARLYRFPYPRADRVMSVCLRVE 181
 Db 122 YSRDGRWRWRHGHKQVLDGNSYDIFLKDLEPIVAVRFRFIPVTDHSMNVCMRVE 181
 QY 182 LYGLWRDGLLSYAPVGTQMYL--SEAYVINDSYDHTVCGLOYGLGLQADGVGLD 239
 Db 182 LYGLWRDGLLSYAPVGTQMYL--SEAYVINDSYDHTVCGLOYGLGLQADGVGLD 240
 QY 240 DFKSKQELRWPCYDYGWNSHFSFGYVEMEFEDRLRAFAQMVHCHNNHHTLGLARLP 299
 Db 241 DFTQTEHYHWPQYDYGWNSHFSFGYVEMEFEDRLRAFAQMVHCHNNHHTLGLARLP 299
 QY 300 GVECRFRGPANAWGEPVHNLGNLGDPRARAVSVPLGGVRARFLQCRFLFAGPWLFF 359
 Db 301 EVQCYF-RSEASEWEPNALSFLVLDVNPASRFTVPLHHRWASAIKQYHAFDWTMMF 359
 QY 360 SEISFISDV-VNNSSPALGTFPPAPWPPGPPPTNFSSLELEPRCQOPVAKPEGSPAI 418
 Db 360 SEITFOSDAAMYNNSEAL-----PTSP-----NAPTYYDPMKLVDDSNTRI 400
 QY 419 LIGCLVAIILLLLLLIALMLRHLWRLLSKAERRVLEELTVHLSVPGDITILNRR--P 476
 Db 401 LIGCLVAIIFILLAIIVILWRQFWOKMLEKASRRMLDDEMTVSLPSLSDSMFNNR 460
 QY 477 GPREP-----PPYQEPFRGNPNPHSAPCVNGSALLSNPAYRLLATYARP 523
 Db 461 SPSEOGSNTSYDRIFPLRPDYQEP-----SRLKLPF----- 495
 QY 524 PRGPGPTTANAKPTNTQAYSGDYMEPEKPGAPILLPPPPQNSVPHYADIVTLQVGTG 583
 Db 495 -----APGEEESGCGVWKPQVSGP-----EGVPHYADIVTLQVGTG 535
 QY 584 NYAVPALPGAVGDGPPRV-DPPRSRLRFLKELGEGQGEVHLCVDSQDLVSLDFPL 642
 Db 536 NYISVAVTMDLLSGKDVAVEEPFKLLTFKELGEGQGEVHLCVDSQDLVSLDFPL 595
 QY 643 NVKRGHPLLVAVKILRPDATKNARNDFLKEVKIMSLKDPNIIIRLLGVCVQDDPLCMITD 702
 Db 596 DVSANQPLVAVKMLRADANKARNDFLKEIKIMSLKDPNIIIRLLGVCVQDDPLCMITE 555
 QY 703 YMENGLDNOFTSAHOLEDKAAEGAPGDGQAAGPTISYPMILHVAQAQASGMRYLATLNF 762
 Db 656 YMENGLDNOFTSAHOLEDKAAEGAPGDGQAAGPTISYPMILHVAQAQASGMRYLATLNF 706
 QY 763 VHRDLATNCLVGENFTIKIADFGMSRLNYAGDYRVQGRVAVLPIRWMAWECILMGKFTT 822
 Db 707 VHRDLATNCLVGENFTIKIADFGMSRLNYAGDYRVQGRVAVLPIRWMAWECILMGKFTT 766

QY 701 TDYMGDLNOLFSAHQLEDKAAEGPDGQAAQC---PTISYPMLLHVAQIASGRYL 757
 Db 613 FEYKHGDLNKLRAH-----GPDVLMWAGNPAELTQSMHLHIAAGNVYL 663
 QY 758 ATLVNVRDLATNCLNGENTIKIADFGMRNLVAGDYRVQGRVLPINWAWECILM 817
 Db 664 ASQHEVHRDLATNCLNGENLLVKGIDFGMSRDVYSTDYRVGGHTMLPIRMMPPESIMY 723
 QY 818 GFTTASDVWAFGVTLWELVLMCLRAQFGOLTDEQVNIENAGEFFRDGQVYLSRPPACP 877
 Db 724 RKTETSDVMSGLVNLWEIFYTKG-OPWYQSNNEVIECI-----TQGR--VLQRPRTCP 775
 QY 878 QGLYELMLRCWSRESEQRPPFSQLHREL 905
 Db 776 KEVYDMLGCWOREPHMLNIKEIHSLL 803
 RESULT 10
 S06943
 brain-derived neurotrophic factor receptor precursor - mouse
 N:Alternate names: receptor tyrosine kinase trkB
 N:Contains: protein-tyrosine kinase (EC 2.7.1.112)
 C:Species: Mus musculus (house mouse)
 C:Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 21-Nov-1997
 C:Accession: S06943
 R:Klein, R.; Parada, L.F.; Coullier, F.; Barbacid, M.
 EMBO J. 8, 3701-3709, 1989
 A:Title: trkB, a novel tyrosine protein kinase receptor expressed during mouse neural de
 A:Reference number: S06943; MUID:90059970
 A:Accession: S06943
 A:Molecule type: mRNA
 A:Residues: 1-821 <KLE>
 A:Cross-references: EMBL:X17647; NID:g55505; PID:g55506
 A:Genetics:
 A:Gene: trkB
 C:Function:
 C:Description: regulation of nervous system development; receptor for brain-derived neur
 C:Superfamily: nerve growth factor receptor, high affinity; leucine-rich alpha-2-glycop
 C:Keywords: alternative splicing; ATP; autophosphorylation; glycoprotein; growth factor
 inase
 F:1-31/Domain: signal sequence #status predicted <SIG>
 F:32-821/Product: brain-derived neurotrophic factor receptor #status predicted <MAT>
 F:32-434/Domain: extracellular #status predicted <EXT>
 F:67-91/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
 F:92-115/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
 F:116-138/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
 F:139-150/Domain: leucine-rich alpha-2-glycoprotein repeat homology #status atypical
 F:435-451/Domain: transmembrane #status predicted <TMN>
 F:452-821/Domain: cytosolic #status predicted <CYT>
 F:535-813/Domain: protein kinase homology <KIN>
 F:543-551/Region: protein kinase ATP-binding motif
 F:67,95,121,178,205,241,254,280,325,338,350,411/Binding site: carbohydrate (Asn) (coval
 F:571/Active site: Lys #status predicted
 F:706/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted
 F:816/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 13.5%; Score 659.5; DB 1; Length 821;
 Best Local Similarity 30.4%; Pred. No. 1.6e-29;
 Matches 201; Conservative 72; Mismatches 188; Indels 201; Gaps 22;
 QY 373 SPALGCTFPAPWPPGPPPTNFSSLELEPRGQOPVAKPEGSPATLIGCLVAILLLL 432
 Db 217 SCVGGDPLTLVMDVNLVSKHNNETSHTQSGSLRITNSSDSDSGKQISCV----- 268
 QY 433 IIALMLRLHWRLLSKAERVLFEELTVHLSV---PGDTILNNRPPGPREP----- 483
 Db 268 -----AENLVGDDQDSVNLTVHFAPTITFL-----ESPTSDHHC 302
 QY 483 -PYOEPGRPNPSPAPCPVNGS-----ALLSNPA-----Y 513
 Db 303 IFF---TVRCNPKPALQWFWNGAILNESKYICTKIHNTHTEYHGCLQDNPTHMNGDY 359

QY 514 RLL-----ATYAPRPGPPPTPAWAKPTNTQAYSGDYMEP--E 551
 Db 360 TLMAKNEKDEROISAHFPMRPGVDYETNPYPEVLYEDWTTPTDI-----GDTINKSNE 415
 QY 552 KFCAPLLPPPPQNSPHYAEADIVTLOGV----- 581
 Db 416 IPSTDVADQSNREHLSVYAVVIVSVVGFCLLVMLLLKLARHSKFGMKGPASVINSDDD 475
 QY 581 -----TGGNTYAVPALPPGAVGDGP---PRVDFP-----RS 608
 Db 476 SASPLHINGSNTFSSSEGGPDVLIWTKIPVIENQYFQITNSQLKPDFTVOHIKRH 535
 QY 609 RLREFEKLGEQFGEVHLCE---VDSQDLVSLDFPLNVKRGHPLLVAVKILRPAATKNA 665
 Db 536 NIVLRELGEAGFKVFLAECYNLCPEQD-----KILVAVKTLK-DASDNA 580
 QY 666 RNDFLKEVKIMSLKDPNLIIRLLGVCVODDPLCLMTIDYMGDLNOLFSAHQLEDKAAE 724
 Db 581 RKDFHREAELLTNLQHEHIVKFGVCEGDPIMVFYMKHGDNLNKLRAHGPDAVLMAE 640
 QY 725 GAPDGOAAQGGPT-ISYPMLLHVAQIASGRYLATLVNVRDLATNCLNGENTIKIA 783
 Db 641 GNP-----PTELTQSMHLHIAQIAAGVYLSAQHFVHRDLATNCLNGENLLVKG 692
 QY 784 DFCMSRLNAGDYRVQGRVLPINWAWECILMGKFTTASDVWAFGVTLWELVLMCLRAQ 843
 Db 693 DFCMSRDVYSTDYRVGGHTMLPIRMMPPESIMYRKTETSDVMSGLVNLWEIFYTKG-Q 751
 QY 844 PFGQLTDEQVNIENAGEFFRDGQVYLSRPPACPQGLYELMLRCWSRESEQRPPFSQLHR 903
 Db 752 PWYQLSNNEVIECI-----TQGR--VLQRPRTCPQVYELMLGCWOREPHTRKNIKSIHT 804
 QY 904 FL 905
 Db 805 LL 806
 RESULT 11
 A39667
 brain-derived neurotrophic factor receptor precursor - rat
 N:Alternate names: receptor tyrosine kinase trkB.FL
 N:Contains: protein-tyrosine kinase (EC 2.7.1.112)
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 27-Nov-1991 #sequence_revision 27-Nov-1991 #text_change 21-Nov-1997
 C:Accession: A39667
 R:Middlemas, D.S.; Lindberg, R.A.; Hunter, T.
 Mol. Cell. Biol. 11, 143-153, 1991
 A:Title: trkB, a neural receptor protein-tyrosine kinase: evidence for a full-length
 A:Reference number: A39667; MUID:91094826
 A:Accession: A39667
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-821 <MID>
 A:Cross-references: GB:M55291; NID:g207473; PID:g207474
 C:Function:
 A:Description: regulation of nervous system development; receptor for brain-derived n
 C:Superfamily: nerve growth factor receptor, high affinity; leucine-rich alpha-2-glyc
 C:Keywords: alternative splicing; ATP; autophosphorylation; glycoprotein; growth fact
 inase
 F:1-31/Domain: signal sequence #status predicted <SIG>
 F:32-821/Product: brain-derived neurotrophic factor receptor #status predicted <MAT>
 F:32-434/Domain: extracellular #status predicted <EXT>
 F:67-91/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
 F:92-115/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
 F:116-138/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
 F:139-150/Domain: leucine-rich alpha-2-glycoprotein repeat homology #status atypical
 F:435-451/Domain: transmembrane #status predicted <TMN>
 F:452-821/Domain: cytosolic #status predicted <CYT>
 F:535-813/Domain: protein kinase homology <KIN>
 F:543-551/Region: protein kinase ATP-binding motif
 F:67,95,121,178,205,241,254,280,325,338,350,411/Binding site: carbohydrate (Asn) (cov
 F:571/Active site: Lys #status predicted
 F:705/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predi

F:816/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 13.5%; Score 659.5; DB 1; Length 821;
Best Local Similarity 30.4%; Pred. No. 1.6e-29;
Matches 201; Conservative 72; Mismatches 188; Indels 201; Gaps 22;

QY 373 SPALGTEPPAPWPGPPPTNFSSLELEPRGOQVAKPEGSPTAILGICLVAILLLL 432
DB 217 SCVSGDPLTLYWDVGNLVSKHMMNETSQTSLRITNISDDSKQISCV----- 268
QY 433 IIALMLRLHWRLLSKARRVLEELTVHLSV---PGDTILNNRPGRPP----- 483
DB 268 -----AENLVGEDQDSVNLTVHFAPITEL-----ESPSTDHWC 302
QY 483 -PYQPRPRGNPHSAPCPVNGS-----ALLSNPA-----Y 513
DB 303 IPF---TVRGNPKPALQFYNGAILNESKYICTKIHTVHTYHGCCLQDNFTHNNGDY 359
QY 514 RLLL-----ATYARPPRGPGPPPTPAWAKPTNTQAYSGDYNEP--E 551
DB 360 TLMKANEYKDKQISAHFMGRPGVDYETNPPEVLYEDWTTPTDI-----GDTNKSNE 415
QY 552 KPGAPLLPPPPONSYPHVAEDIVTQGV-----KILVAVTKL-DASDNA 580
DB 416 IPSTVDADQTNREHLSVAVVVIASVGVFCLLVMLLLKLARHSGKMGKSPASVISNDDD 475
QY 581 -----TGNTYAVPALPGAVGDP---PRVDFP-----RS 608
DB 476 SASPLHHISNGSNTSPSEGGPDVILGWTKIPVIENQYFGIINSQKLPFTFVQHVKRH 535
QY 609 RLRFKEKLGEGOFGEVHICE---VDSQDLVSLDFPLNVRKHPHLLVAVKILRPDATKNA 665
DB 536 NIVLRKELGEGAFKVFLEACYNLCPEQD-----KILVAVTKL-DASDNA 580
QY 666 RNDFLKEVKIKMSRLKDPNIIRLLGVCVQDDPLCMITDYMENGLNQFLSAHQLED-KAAE 724
DB 581 RKDFHREALLTNLQHEHIVFYGVCGVDEPLIMVFEYMKHGDNLKFLRAHGPDAVLMAE 640
QY 725 GAPDGOAAQOQPT-ISTYPMLLHVAQAASGRMYLATLNFVHRDLATRNCLVGENFTIKIA 783
DB 641 GNP-----PTELTQSMHLIAQQAAGVMYLASQHFVHRDLATRNCLVGENLLVKIG 692
QY 784 DFGMSRLNLYAGDYRYVQGRAVLPTRMAWECILMGKFTTASDVAFGVTWLWEVLMCLRAQ 843
DB 693 DFGMSRDVYTDYRYVGGHTMLPIRMPPESIMYRKFTTESDVWSGLVGLWEIYFTYK-Q 751
QY 844 PFGQLTDEQVIENAGEFFRDQGRVYLSRPPACPGQVYELMLRCWSESRQRPFPFSLHR 903
DB 752 PWYQLSNNEVIECI-----TQGR--VLQRPRTCPQEVYELMLGCGWQREPHTRNKNIHT 804
QY 904 FL 905
DB 805 LL 806

RESULT 12

A56853
brain-derived neurotrophic factor receptor precursor - human
N:Alternate names: receptor tyrosine kinase trkb
N:Contains: protein-tyrosine kinase (EC 2.7.1.112)
C:Species: Homo sapiens (man)
C:Date: 11-Aug-1995 #sequence_revision 11-Aug-1995 #text_change 21-Nov-1997
C:Accession: A56853; I56557
R:Nakagawara, A.; Liu, X.G.;
Genomics 25, 538-546, 1995
A:Title: Cloning and chromosomal localization of the human TRK-B tyrosine kinase receptor
A:Reference number: A56853; MUID:95309922
A:Accession: A56853
A:Molecule type: mRNA
A:Residues: 1-822 <NAK>
A:Cross-references: GB:U12140; NID:9525313; PID:g530791
R:Shelton, D.L.; Sutherland, J.; Gripp, J.; Camerato, T.; Armanini, M.P.; Phillips, H.S.

J. Neurosci. 15, 477-491, 1995
A:Title: Human trks: molecular cloning, tissue distribution, and expression of extrac
A:Reference number: I56557; MUID:95123473
A:Accession: I56557
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-822 <SHE>
A:Cross-references: GB:S76473; NID:9913717; PID:9913718
C:Genetics:
A:Gene: GDB:NTRK2; trkb
A:Cross-references: GDB:127898; OMIM:600456
A:Map position: 9q22.1-9q22.1
C:Function:
A:Description: regulation of nervous system development; receptor for brain-derived n
C:Superfamily: nerve growth factor receptor, high affinity; leucine-rich alpha-2-glyc
C:Keywords: alternative splicing; ATP; autophosphorylation; glycoprotein; growth fact
inase

F:1-31/Domain: signal sequence #status predicted <SIG>
F:32-822/Product: brain-derived neurotrophic factor receptor #status predicted <MAT>
F:32-435/Domain: extracellular #status predicted <EXT>
F:97-91/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
F:92-115/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F:116-138/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F:139-150/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F:436-452/Domain: transmembrane #status predicted <TMN>
F:453-822/Domain: cytosolic #status predicted <CYT>
F:536-814/Domain: protein kinase homology <KIN>
F:544-552/Region: protein kinase ATP-binding motif
F:57,95,121,178,204,280,325,338,350,412/Binding site: carboxyhydrate (Asn) (cov
F:572/Active site: Lys #status predicted
F:706/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predi
F:817/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 13.4%; Score 658; DB 1; Length 822;

Best Local Similarity 47.4%; Pred. No. 2e-29;

Matches 144; Conservative 45; Mismatches 79; Indels 36; Gaps 9;

QY 607 RSLRFEKLGEGOFGEVHICE---VDSQDLVSLDFPLNVRKHPHLLVAVKILRPDATK 663
DB 535 RHNVILKRELGEAGFGKVFLEACYNLCPEQD-----KILVAVTKL-DASD 579
QY 664 NARNDLKEVKIKMSRLKDPNIIRLLGVCVQDDPLCMITDYMENGLNQFLSAHQLED-KA 722
DB 580 NARKDPHREALLTNLQHEHIVFYGVCGVDEPLIMVFEYMKHGDNLKFLRAHGPDAVL 639
QY 723 AEGAPDGOAAQOQPT-ISTYPMLLHVAQAASGRMYLATLNFVHRDLATRNCLVGENFTIK 781
DB 640 AEGNP-----PTELTQSMHLIAQQAAGVMYLASQHFVHRDLATRNCLVGENLLVK 691
QY 782 IADFGMSRLNLYAGDYRYVQGRAVLPTRMAWECILMGKFTTASDVAFGVTWLWEVLMCLR 841
DB 692 IADFGMSRDVYTDYRYVGGHTMLPIRMPPESIMYRKFTTESDVWSGLVGLWEIYFTYK 751
QY 842 AQPFGQLTDEQVIENAGEFFRDQGRVYLSRPPACPGQVYELMLRCWSESRQRPFPFSQL 901
DB 752 -QPWYQLSNNEVIECI-----TQGR--VLQRPRTCPQEVYELMLGCGWQREPHTRNKNIKI 803
QY 902 HREL 905
DB 804 HTLL 807

RESULT 13

TRVTTB
nerve growth factor receptor precursor, high affinity - black rat
N:Alternate names: receptor tyrosine kinase trka
N:Contains: protein-tyrosine kinase (EC 2.7.1.112)
C:Species: Rattus rattus (black rat, roof rat)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 21-Nov-1997
C:Accession: A41981
R:Weakin, S.O.; Suter, U.; Drinkwater, C.C.; Welcher, A.A.; Shooter, E.M.
Proc. Natl. Acad. Sci. U.S.A. 89, 2374-2378, 1992

A:Title: The rat trk protooncogene product exhibits properties characteristic of the sld
A:Reference number: A41981; MUID:92196121
A:Accession: A41981
A:Molecule type: mRNA
A:Residues: 1-799 <MEA>
A:Cross-references: GB:M85214; NID:g207481; PID:g207482
A:Note: sequence extracted from NCBI backbone (NCBI:88433, NCBI:88434)
C:Comment: The proto-oncogene trka is activated by gene fusion. The amino end of several
C:Function:
A:Description: regulation of nervous system development; receptor for nerve growth factor
C:Superfamily: nerve growth factor receptor, high affinity; leucine-rich alpha-2-glycoprotein
C:Keywords: ATP; autophosphorylation; glycoprotein; growth factor receptor; phosphoprotein
C:Specific protein kinase
F:1-32/Domain: signal sequence #status predicted <SIG>
F:33-799/Product: nerve growth factor receptor, high-affinity #status predicted <MAT>
F:33-424/Domain: extracellular #status predicted <EXT>
F:67-91/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
F:92-115/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F:116-138/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F:139-150/Domain: leucine-rich alpha-2-glycoprotein repeat homology #status atypical <LR>
F:425-441/Domain: transmembrane #status predicted <TMN>
F:442-799/Domain: cytosolic #status predicted <CYT>
F:511-791/Domain: protein kinase homology <KIN>
F:519-527/Region: protein kinase ATP-binding motif
F:67,121,190,204,255,264,320,325,341,361,404/Binding site: carbohydrate (Asn) (covalent)
F:547/Active site: Lys #status predicted
F:683/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted
F:794/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 13.3%; Score 653.5; DB 1; Length 799;
Best Local Similarity 29.8%; Pred. No. 3,4e-29;
Matches 190; Conservative 78; Mismatches 155; Indels 215; Gaps 21;

QY 335 SVPLGGRVARELOC-----RFLFAGPWLLESEISFI----- 366
DB 289 SVHLGKAEQHHWCIPFSDGQAPSLRWFFNGS--VLNLSFTIFLSEATNETMRH 346
QY 366 -----SDVYNNSSPALGGTFP-----PAPWMPGPPPTNFSLELEPR 403
DB 347 GCLRLNQTHVNGNYTLAANPYGQAASTMAAFMDNPFENFEDPIPVSEFVDNST 406
QY 404 GOQPVAKPEGSPATLIGCLVAITLLILLIITALLMLRLHWRLLSKAERR-----V 454
DB 407 SRDPEKKDETDFGVSAGVLAASALFL-SALLL-----VLNKGQSKFGINRAV 458
QY 455 LEE-----LTVHLSVPGTILNNRPGPREPPYQEPGRNPPHSPACVPNGSAL--L 507
DB 459 LAPEDGLAMSLHFMTLGGSSL-----SPTGKSGSLGQHI 493
QY 508 LSNPAYRLLATYARPPRGPGPTPAWAKPTNTQAYSGDYMEPEKPGAPLLPPPPQNSVP 567
DB 494 MENPOY-----FS-----DTCVH 506
QY 568 HYAEADIVTLOGVTGGNTYAVPALPPGAVGDGPPRDLRPRRLRKEKLGQGEVHLC 627
DB 507 HTRQDII-----LKWELGSGAFGKFLA 530
QY 628 E---VDSQDLVDFPLNVRKGPLLVAVKTLRPDATKNARNDLKEVKIMSLKDPNI 684
DB 531 ECYNLLNDD-----KMLVAVKALK--ETSENARQDFHRAEELLTLQHQHI 575
QY 685 IRLGLVCVQDDPLCMITDYMENDLNQFLSAHQLEDKAAEGAPDQGAQOQTSYPMLL 744
DB 576 VRFEGVCTEGGLLVAVFMYMRHGLDLNRLRSHGPDAKLAGGE---DYAPGP-LGLGOLL 631
QY 745 HVAQAQASGRVYLATLNFVHRDLATRNCLVGENFTIKIADFGMSRNLVYAGDYRQGRV 804
DB 632 AVASQVAAQVYLAHLFVHRDLATRNCLVGLVGVKIDFGMSRDIYSTDYRVGRTM 691
QY 805 LPIRMWAEICLMGKFTTASDVAFGVTLMVEVLMCRAPQGLTDEQVIENAGEFFRDQ 864
DB 692 LPIRMWPPESILYRKFTESDWSFGVVLWEIFTYTKG--QPMFQLSNTEAIECI-----TQ 745

QY 865 GRQVYLSPRPACPOGLXELMLRCSRESEQRPPFSQLH 902
DB 746 GRE--LERPRCPDPVYAIMRGWCQREPOQRSLMKDVH 781

RESULT 14
A51178
A51178
neurotrophin receptor trkC precursor - human
C:Species: Homo sapiens (man)
C:Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 24-Sep-1998
C:Accession: A51178
R:McGregor, L.M.; Baylin, S.B.; Griffin, C.A.; Hawkins, A.L.; Nelkin, B.D.
Genomics 22, 267-272, 1994
A:Title: Molecular cloning of the cDNA for human TrkC (NTRK3), chromosomal assignment
A:Reference number: A51178; MUID:95104834
A:Accession: A51178
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-825 <MCG>
A:Cross-references: GB:U05012; NID:g442389; PID:g442390
C:Genetics:
A:Gene: GDB:NTRK3
A:Cross-references: GDB:127899; OMIM:191316
A:Map position: 13q24-15q25
C:Superfamily: nerve growth factor receptor, high affinity; leucine-rich alpha-2-glyc
C:Keywords: alternative splicing; ATP; growth factor receptor
F:79-103/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
F:104-127/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F:128-150/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F:151-162/Domain: leucine-rich alpha-2-glycoprotein repeat homology #status atypical
F:536-817/Domain: protein kinase homology <KIN>
F:544-552/Region: protein kinase ATP-binding motif

Query Match 13.3%; Score 651.5; DB 2; Length 825;
Best Local Similarity 32.9%; Pred. No. 4.6e-29;
Matches 174; Conservative 66; Mismatches 132; Indels 157; Gaps 18;

QY 490 RGNPPHSPACVPNG-----SALLSNPA-----YRLLIATYAR 522
DB 326 RGNPPPTLHLWNGOPLRESKIIHVEYQEGEISGCLLFNKPHTYNNNGVTLI---AK 381
QY 523 PPRGPGPPPAWAKTNTQAYSGDYME---PEKGPALL---PPPPNSVPHYAEADIV 575
DB 382 NPLG-----TANQINGHELKEPPESTONFILFDEVSPPTITVTHKPEEDTF 430
QY 576 TLQGYTGGNTYAVPAL-----P 592
DB 431 GVSIAVGLAAAFACVLLVLFVMIKYGRRKFGMKGPVAVISGEDSDASPLHHIHGITT 490
QY 593 PGAVGDGP-----PRVDFP-----RSRLRFKEKLGEGQFG 622
DB 491 PSSLDAGPTVVTGTRIPVNIENPOYFROGHNCHKPDYVQHIKRRDVLKRELGEAGF 550
QY 623 EYHLCVDSQDLVSLDFPLNVRKGPLLVAVKILRPDATKNARNDLKEVKIMSLKDP 682
DB 551 KVFLAEC-----YNLSPTK-DKMLVAVKALK-DPTLAARKDFQREAEELLNLQHE 598
QY 683 NIIRLLGVCVQDDPLCMITDYMENDLNQFLSAHQLEDKAAEGAPG---DQAAQGP- 737
DB 599 HIVKFGVCGDGDPLIMVFEYMKHGDNLKFLAH-----GNAMTLVDGQPPQAKG 649
QY 737 TISYPMLLHVAQAQASGRVYLATLNFVHRDLATRNCLVGENFTIKIADFGMSRNLVYAGDY 796
DB 650 ELGLSQMLHIAQIASGRVYLAHQHEVHRDLATRNCLVGNLVLKIGDGMRSRDVSTDY 709
QY 797 YRVQGRAVLPIRMWAEICLMGKFTTASDVAFGVTLMVEVLMCRAPQGLTDSQVTEN 856
DB 710 YRVGGHTMLPIRMWPPESILYRKFTESDWSFGVVLWEIFTYTKG--QPMFQLSNTEVIEC 768
QY 857 AGEFFRDQGRVYLSRPPACPOGLXELMLRCSRESEQRPPFSQLHREL 905

Db 769 I-----TQGR--VLERPRVCPKEVDVLMGLCWQREPOQRLNIKEIKIL 810

RESULT 15

A40026

neurotrophin-3 receptor precursor - pig

N:Alternate names: receptor tyrosine kinase trkc

N:Contains: protein-tyrosine kinase (EC 2.7.1.112)

C:Species: Sus scrofa domestica (domestic pig)

C>Date: 17-Jan-1992 #sequence_revision 17-Jan-1992 #text_change 21-Nov-1997

C:Accession: A40026

R:Lamballe, F.; Klein, R.; Barbacid, M.

Cell 66, 967-979, 1991

A:Title: trkc, a new member of the trk family of tyrosine protein kinases, is a receptor

A:Reference number: A40026; MUID:91364178

A:Accession: A40026

A:Molecule type: mRNA

A:Residues: 1-825 <LAW>

A:Cross-references: GB:M80800; NID:g164698; PID:g164699

A:Experimental source: adult brain

C:Function:

A:Description: regulation of nervous system development; receptor for neurotrophin-3

C:Superfamily: nerve growth factor receptor, high affinity; leucine-rich alpha-2-glycop

C:Keywords: ATP; autophosphorylation; brain; glycoprotein; growth factor receptor; phosph

F:1-31/Domain: signal sequence #status predicted <Sig>

F:32-825/Product: neurotrophin-3 receptor #status predicted <NAT>

F:32-436/Domain: extracellular #status predicted <EXT>

F:79-103/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>

F:104-127/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>

F:128-150/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>

F:151-162/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>

F:437-453/Domain: transmembrane #status predicted <TMN>

F:454-825/Domain: cytosolic #status predicted <CYT>

F:536-817/Domain: protein kinase homology <KIN>

F:544-552/Region: protein kinase ATP-binding motif

F:58,72,79,133,163,203,218,232,259,267,272,294,375,388/Binding site: carbohydrate (Asn)

F:572/Active site: Lys #status predicted

F:709/Binding site: phosphate (Tyr) (covalent) #status predicted

F:820/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 13.1%; Score 642.5; DB 1; Length 825;

Best Local Similarity 30.4%; Pred. No. 1.4e-28;

Matches 190; Conservative 82; Mismatches 179; Indels 175; Gaps 26;

QY 309 PAMAW--EGEPMRHLGNLGDPRARAVSVPLGGRVAREFLQCRFLFAGFWLFLFSIFIS 366

Db 331 PTLHLHNGOPLR-----ESKITHVEYQGEVSE--GC-----LLFNKPTHYN 372

QY 367 D---VYNNSSP-----ALGGTFPPAPWPPGPPPT-NFSSL-ELEPRGQOPVA-KPEGGS 414

Db 373 NGNYTLNROEPLGTANTQINGHFLKEPF---PESTDNFVSFEVSPPTITVTHKPEED 428

QY 415 PTAILGICVAIILLILLIALLMLRWRLLSKAERRVLEELTVHLSVPGDTILINN 474

Db 429 TFGVSTAVGLAFACVLLVLLFIMINKYGR--SK-----FGMKGPVAVISG 473

QY 475 RQGPPEPPPYQEP-----RPRGNPPHSAPCPVNGSALLSNPAYRLLLATYARPPR 525

Db 474 EEDSASPLHHDQPHHHHTLITCRAGHSVIGMTRIP-----VIENPOY-----517

QY 526 GPGPPTPAKPTNTQAYSGDMEPEKPGAPLLPPPPQNSVPHYAEADIVTLQGVGTGNT 585

Db 517 -----FROGHNCH-----KPDITYVQHKRRDIV-----540

QY 586 YAVPALPGAVGDGPRVDFPRSLRFEKLGEGFVHLCFVDSQDLYSLDFPLNVR 645

Db 540 -----LKRELGEAGFVKVFLAECYNLS-----PTKV 566

QY 646 KGHPLLVAVKILRPDATKNARNDLFKEVKIMSKLDPNIRLLGVCVODDPLCMITDYME 705

Db 567 ----MLVAVKALK-DPTLAARKDFQREAEILLNLOHEHIVKFYGVCGDGDPLIMVFEYMK 621

QY 706 NGDINQFLSAHQLEDKAAEGAPG-----DQAAAGQP--TISYPMLLHVAAQIASGRMYLAT 759

Db 622 HGDNLKFLRAH-----GPDAMILVDGQPRQANGELGSLQMLHIASQICSGWYLAS 672

QY 760 LNFVHRDLATRNCLVGENFTFIKADFGMSRNLYAGDYRVQGRAVLPIRWNWECILMGK 819

Db 673 QHFVHRDLATRNCLVGANLLVKIGDFGMSRDVYSTDYRVGGHTMLPIRWNWPPESIMYRK 732

QY 820 FTTASDYWATGVTILWEVLMCLCRAQPFQGLTDEQVIENAGFEFFRDQGRVYLSRPPACPOG 879

Db 733 FTTESDYWSTGVTILWEVLTFTYTKG-QPWQLSNTVEIECI-----TQGR--VLERPRVCPKE 784

QY 880 LYELMLRCWSESEQRPPFSQLHREL 905

Db 785 VIDVMLGCWQREPOQRLNIKEIKIL 810

Search completed: November 4, 1999, 01:52:29
Job time: 7584 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 4, 1999, 05:43:04 ; Search time 752.95 Seconds
(without alignments)
9527.996 Million cell updates/sec

Title: US-08-170-558-3
Perfect score: 3637
Sequence: 1 GAATTCGTGAGTCGACGTTG.....AAAAAAAAAAAAAAAAAAAA 3637

Scoring table: IDENTITY_NUC
Searched: 2546578 seqs, 986266752 residues

Database :
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*
44: gb_est25:*
45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*

54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
C 1	638	17.5	652	49	AL047539	AL047539 DKF2p586C
C 2	572.4	15.7	772	54	HSN008101	AL043251 Homo sapi
C 3	505.4	13.9	507	45	A1342378	A1342378 qt27e10.x
C 4	493.4	13.6	495	54	HSN010070	A1045220 Homo sapi
C 5	475.8	13.1	481	38	AA781458	AA781458 aj26f12.s
C 6	471.6	13.0	519	31	AA325140	AA325140 EST28097
C 7	470.8	12.9	474	43	A1204173	A1204173 qd70e12.x
C 8	457.6	12.6	464	45	A1365945	A1365945 ao91a04.x
C 9	448.4	12.3	532	28	AA116659	AA116659 mq23f05.r
C 10	444.4	12.2	446	44	A1275058	A1275058 ql72d06.x
C 11	440.4	12.1	453	39	AA831559	AA831559 Oc84d05.s
C 12	432.2	11.9	542	40	AA944928	AA944928 EST200427
C 13	416.4	11.4	429	33	AA430320	AA430320 zw07f11.r
C 14	416.2	11.4	422	44	A1280882	A1280882 qk47b03.x
C 15	416	11.4	465	27	C04008	C04008 C04008 Huma
C 16	411.4	11.3	524	42	A1145187	A1145187 UI-R-BT0-
C 17	411.2	11.3	416	48	A1588864	A1588864 tq28d12.x
C 18	408.4	11.2	418	20	T09338	T09338 EST07231 in
C 19	406.4	11.2	408	48	A1567223	A1567223 tp51e07.x
C 20	400.8	11.0	404	43	A1214511	A1214511 qm27g07.x
C 21	396.6	10.9	422	34	AA646452	AA646452 zx85c05.r
C 22	393	10.8	393	47	A1494466	A1494466 qz16a09.x
C 23	392.4	10.8	406	50	A1686734	A1686734 tx09d08.x
C 24	391.8	10.8	406	40	AA938929	AA938929 op83c09.s
C 25	389.4	10.7	505	44	A1323681	A1323681 mq23f05.x
C 26	381	10.5	477	46	A1430426	A1430426 mf77g10.y
C 27	380.4	10.5	382	38	AA808281	AA808281 oc41d07.s
C 28	379.8	10.4	383	20	T09337	T09337 EST07230 in
C 29	378.4	10.4	404	50	A1689485	A1689485 tx94g05.x
C 30	370.2	10.2	440	45	A1326485	A1326485 mq23f05.y
C 31	366.4	10.1	368	38	AA746483	AA746483 nw62e12.s
C 32	364.4	10.0	413	34	AA646451	AA646451 zx85c05.s
C 33	363.8	10.0	378	47	A1476072	A1476072 tl97e09.x
C 34	360.6	9.9	571	20	X73728	X73728 CHESTW025 G
C 35	355.8	9.8	433	46	AA956276	AA956276 UI-R-E1-f
C 36	349.4	9.6	383	33	AA385987	AA385987 EST99705
C 37	348.8	9.6	425	51	A1712520	A1712520 UI-R-AFI-
C 38	347.2	9.5	364	36	AA612719	AA612719 ng26c10.s
C 39	346.8	9.5	361	35	AA574033	AA574033 nkl1a11.s
C 40	344.4	9.4	347	50	A1686242	A1686242 tu40h04.x
C 41	343	9.4	424	32	AA338319	AA338319 EST43238
C 42	335.8	9.2	339	49	A1620626	A1620626 tu47g12.x
C 43	334.8	9.2	730	26	W22881	W22881 75810 Human
C 44	332.4	9.1	422	26	W98395	W98395 mg20b06.r1
C 45	330	9.1	379	20	T29903	T29903 EST99706 Hu

ALIGNMENTS

RESULT 1
AL047539/c
LOCUS AL047539 652 bp mRNA EST
DEFINITION DKF2p586C0921.sl 586 (synonym: hute1) Homo sapiens CDNA clone
ACCESSION AL047539
NID 94728535
VERSION AL047539.1 GI:4728535

KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 652)
AUTHORS Ottenwaelder B., Obermaier, B., Mewes, W., Gassenhuber, J. and Wiemann, S.
TITLE EST (Ottenwaelder, et al.)
JOURNAL Unpublished (1999)
COMMENT On Jun 5, 1998 this sequence version replaced gi:3189690.

Contact: Ottenwaelder B
MIPS
Am Klopferspitz 18a D-82152 Martinsried, Germany.
location/Qualifiers
1. 652
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp386C0921"
/clone_lib="586 (synonym: hutel)"
/tissue_type="uterus"
/dev_stage="adult"
/lab_host="DH10B"
/note="vector: pSport1; Site_1: NotI; Site_2: SalI/MLuI"
BASE COUNT 161 a 165 c 174 g 152 t
ORIGIN

Query Match 17.5%; Score 638; DB 49; Length 652;
Best Local Similarity 99.8%; Pred. No. 4.6e-117;
Matches 649; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 2959 AGAAGCCCTCTGCGCCACCAGCTGGTCTGTGGATGGATCTCTCCACCCCTCTCTA 3018
DB 649 AGAAGCCCTCTGCGCCACCAGCTGGTCTGTGGATGGATCTCTCCACCCCTCTCTA 590
QY 3019 GCCATCCTCTGGGAAGGGTGGGAGAATATAGATAGACATGGACATGGCCCATGG 3078
DB 589 GCCATCCTCTGGGAAGGGTGGGAGAATATAGATAGAC-CTGGACATGGCCCATGG 531
QY 3079 AGCACCCTGGGCCCACTGACAACTGATTCCTGGAGAGTGGCTGGCCCGCCAGCTTC 3138
DB 530 AGCACCCTGGGCCCACTGACAACTGATTCCTGGAGAGTGGCTGGCCCGCCAGCTTC 471
QY 3139 TCTCTCCTGTGCACACTGGACCCCACTGGCTGAGATCTGGGGTGGAGGACACAAGA 3198
DB 470 TCTCTCCTGTGCACACTGGACCCCACTGGCTGAGATCTGGGGTGGAGGACACAAGA 411
QY 3199 AGGAGAGAAATGTTTCTTGTGCTGCTCTGTACTTGTCTCAGCTTGGGGTCTTCTC 3258
DB 410 AGGAGAGAAATGTTTCTTGTGCTGCTCTGTACTTGTCTCAGCTTGGGGTCTTCTC 351.
QY 3259 CTCCTCCATCACTGAACACTGGACCTGGGGGTAGCCCGCCCGCCCTCAGTCACCC 3318
DB 350 CTCCTCCATCACTGAACACTGGACCTGGGGGTAGCCCGCCCGCCCTCAGTCACCC 291
QY 3319 CCACCTCCCACTGCAGCTTGTAGTAGAACTTCTTAAGCCCTATACGTTCTTGTGGAG 3378
DB 290 CCACCTCCCACTGCAGCTTGTAGTAGAACTTCTTAAGCCCTATACGTTCTTGTGGAG 231
QY 3379 TAAATATTGGGATGGGGGAAAGAGGAGCAACGGCCCATAGCTTGGGGTGGACATC 3438
DB 230 TAAATATTGGGATGGGGGAAAGAGGAGCAACGGCCCATAGCTTGGGGTGGACATC 171
QY 3439 TCTAGTGTAGCTGCCACATGATTTTCTATAATCACTTGGGGTGTGTACATTTTGGGG 3498
DB 170 TCTAGTGTAGCTGCCACATGATTTTCTATAATCACTTGGGGTGTGTACATTTTGGGG 111
QY 3499 GGAGAGACAGATTTTACACTAATATATGAGACCTAGCTTGGAGCAATTTAATCCCT 3558
DB 110 GGAGAGACAGATTTTACACTAATATATGAGACCTAGCTTGGAGCAATTTAATCCCT 51

QY 3559 GCACCTAGGCAGGTAATAAAGTTGAGTTTTCACAAAAA 3608
DB 50 GCACCTAGGCAGGTAATAAAGTTGAGTTTTCACAAAAA 1

RESULT 2
ID HSM008101
XX standard; RNA; EST; 772 BP.
AC AL043251;
XX
SV AL043251.1
XX
NI el1401623
XX
DT 12-MAR-1999 (Rel. 59, Created)
DT 12-MAR-1999 (Rel. 59, Last updated, Version 1)
XX
DE Homo sapiens mRNA; EST DKFZp434K0223_r1 (from clone
XX DKFZp434K0223)
DE
KW EST; expressed sequence tag.
XX
XX Homo sapiens (human)
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria;
OC Primates; Catarrhini; Hominidae; Homo.
XX
XX [1]
RN 1-772
RA Blum H., Bauersachs S., Mewes W., Gassenhuber J., Wiemann S.;
RT ;
RL Submitted (12-MAR-1999) to the EMBL/GenBank/DBJ databases.
RL MIPS, Am Klopferspitz 18a D-82152 Martinsried, GERMANY
XX
CC Clone from S. Wiemann, sequenced by LMU within the cDNA
CC sequencing consortium of the German Genome Project
CC No si sequence available
CC This clone is available at the RZPD in Berlin
CC Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
CC Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de
XX
XX Key Location/Qualifiers
FH source 1..772
FH /db_xref="taxon:9606"
FH /organism="Homo sapiens"
FH /clone="DKFZp434K0223"
FH /clone_lib="434 (synonym: htes3). Vector pSport1; host
FH DH10B; sites NotI + SalI"
FH /dev_stage="adult"
FH /tissue_type="testis"
XX
SQ Sequence 772 BP; 118 A; 250 C; 260 G; 143 T; 1 other;

Query Match 15.7%; Score 572.4; DB 54; Length 772;
Best Local Similarity 99.3%; Pred. No. 5e-104;
Matches 595; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 38 AGAGATGCTGCCCCCACCCTTAGGCCGAGGGATCAGGAGCTATGGACACAGGCC 97
DB 158 AGAGATGCTGCCCCCACCCTTAGGCCGAGGGATCAGGAGCTATGGACACAGGCC 217
QY 98 TGTCTATCTTACTGCTGCTCTTGGTGGCAAGTGGAGATGCTGACATGAAGGACATT 157
DB 218 TGTCTATCTTACTGCTGCTCTTGGTGGCAAGTGGAGATGCTGACATGAAGGACATT 277
QY 158 TTGATCTCTGCCAAGTGGCCGCTATGCCCTGGGCATGACGACCGGACCATCCACAGACGTG 217
DB 278 TTGATCTCTGCCAAGTGGCCGCTATGCCCTGGGCATGACGACCGGACCATCCACAGACGTG 337
QY 218 ACATCTCTGCTTCCAGCTTCTTGGTTCAGATTCCTACTGCGCCGCCGACAGAGGTGGAGA 277


```
XX CC Clone from S. Wiemann, sequenced by DKFZ within the cDNA
CC CC sequencing consortium of the German Genome Project
CC CC No SI sequence available at the RZPD in Berlin
CC CC This clone is available at the RZPD in Berlin
CC CC Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
CC CC Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de
XX XX Key Location/Qualifiers
FH Key 1. 495
FH source
FH /db_xref="taxon:9606"
FH /organism="Homo sapiens"
FH /clone="DKFZp34H0150"
FH /clone_lib="434 (synonym: htes3). Vector pSport1; host
FH DH10B; sites NotI + SalI"
FH /dev_stage="adult"
FH /tissue_type="testis"
XX XX Sequence 495 BP; 105 A; 145 C; 142 G; 103 T; 0 other;

Query Match 13.6%; Score 493.4; DB 54; Length 495;
Best Local Similarity 99.8%; Pred. No. 2e-88;
Matches 494; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1716 GGACTATATGAGCTGAGAGCCAGGCGCCCGCTTCTGCCCCACGCTCCCCAGAACAG 1775
DB 1 GGACTATATGAGCTGAGAGCCAGGCGCCCGCTTCTGCCCCACGCTCCCCAGAACAG 60
QY 1776 GCTCCCCATTATGCGAGGCTGACATTTTACCTGCGAGGCGCTCACCGGGGCAACAC 1835
DB 61 GCTCCCCATTATGCGAGGCTGACATTTTACCTGCGAGGCGCTCACCGGGGCAACAC 120
QY 1836 CTATCTGTGCTGACATGCCCCCAGGCGGAGTGGGCGGCGGCGGCGGCGGCGG 1895
DB 121 CTATCTGTGCTGACATGCCCCCAGGCGGAGTGGGCGGCGGCGGCGGCGGCGG 180
QY 1896 CCTCGATCTGACATCCGCTTCAAGGAGAAGCTTGGCGAGGCGGCGGCGGAGTGCA 1955
DB 181 CCTCGATCTGACATCCGCTTCAAGGAGAAGCTTGGCGAGGCGGCGGCGGAGTGCA 240
QY 1956 CTTGTGTGAGGTGACAGCCCTCAAGATCTTGCTGAGTCTTGATTTCCCGCTTAATGTGG 2015
DB 241 CTTGTGTGAGGTGACAGCCCTCAAGATCTTGCTGAGTCTTGATTTCCCGCTTAATGTGG 300
QY 2016 TAAGGACACCTTTTGGCTGAGCTGTCAAGATCTTACGCGCAGATGCCCAAGATGC 2075
DB 301 TAAGGACACCTTTTGGCTGAGCTGTCAAGATCTTACGCGCAGATGCCCAAGATGC 360
QY 2076 CAGGAATGATTTTCTTGAAGAGGTGAAGATCATGTGCGAGGCTCAAGGACCCAAACATCAT 2135
DB 361 CAGGAATGATTTTCTTGAAGAGGTGAAGATCATGTGCGAGGCTCAAGGACCCAAACATCAT 420
QY 2136 TCGGCTGTGGCGGTGTGTGTCAGAGAGACCCCGCTCTGATGATTAAGTACTGACTACATGGA 2195
DB 421 TCGGCTGTGGCGGTGTGTGTCAGAGAGACCCCGCTCTGATGATTAAGTACTGACTACATGGA 480
QY 2196 GAACGCGACCTCAA 2210
DB 481 GAACGCGACCTCAA 495

RESULT 5
AA781458/c
LOCUS AA781458 481 bp mRNA EST 31-DEC-1998
DEFINITION aJ26f12.s1 Soares_testis_NHT Homo sapiens cDNA clone 1391471 3'
mRNA sequence.
ACCESSION AA781458
NID 92840789
VERSION AA781458.1 GI:2840789
KEYWORDS EST.
SOURCE human.
```

```
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 481)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1404759.

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 2285 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 480.
Location/Qualifiers
1. 481
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="1391471"
/clone_lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/vector="pVector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech
Laboratories, Inc., and primed with a Not I - oligo(dT)
primer [5'
TGTTACCAATCTGAAGTGGGAGCGCGCCCAATTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo.
BASE COUNT 131 a 117 c 120 g 113 t
ORIGIN

Query Match 13.1%; Score 475.8; DB 38; Length 481;
Best Local Similarity 99.6%; Pred. No. 6.3e-85;
Matches 477; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3124 TGGCCCCCAGCTTCTCTCCCTGTACACACTGGACCCCACTGGCTGAGAATCTGGGG 3183
DB 479 TGGCCCCCAGCTTCTCTCCCTGTACACACTGGACCCCACTGGCTGAGAATCTGGGG 420
QY 3184 GTGAGGAGACAGAGAGAGAAATGTTCTTGTGCTGCTCTGCTCTGCTCTGCTCTC 3243
DB 419 GTGAGGAGACAGAGAGAGAAATGTTCTTGTGCTGCTCTGCTCTGCTCTGCTCTC 360
QY 3244 AGCTGGGCTTCTTCTCTCCATCACCTGAAACACTGGGCTAGCCGCCGCC 3303
DB 359 AGCTGGGCTTCTTCTCTCCATCACCTGAAACACTGGGCTAGCCGCCGCC 300
QY 3304 AGCCCTCAGTACCCCACTTCCCACTGCAGTCTTGTAGCTAGAACTTCTCTAAGCCTA 3363
DB 299 AGCCCTCAGTACCCCACTTCCCACTGCAGTCTTGTAGCTAGAACTTCTCTAAGCCTA 240
QY 3364 TAGCTTCTGTGAGTAATAATTGGGATTTGGGGGAAAGAGGAGCAAGCCCATAGCC 3423
DB 239 TAGCTTCTGTGAGTAATAATTGGGATTTGGGGGAAAGAGGAGCAAGCGCTCATAGCC 180
QY 3424 TTGGGGTTGGACATCTCTAGTGTAGCTGGCACATTGATTTTCTATAATCACTTGGGGTT 3483
```

```

Db 179 TTGGGTTGGACATCTCTAGTGTAGCTGCCACATGATTTTCTTATAATCACTTGGGTT 120
QY 3484 TGTACATTTTGGGGGAGAGACACAGATATTTTACACTAATATATGGACCTAGCTTTGAGG 3543
Db 119 TGTACATTTTGGGGGAGAGACACAGATTTTACACTAATATATGGACCTAGCTTTGAGG 60
QY 3544 CAATTTTAAATCCCTGCACATAGCAGGTAATTAAGTTTGGATTTTCCACAAAAA 3602
Db 59 CAATTTTAAATCCCTGCACATAGCAGGTAATTAAGTTTGGATTTTCCACAAAAA 1

RESULT 6
AA325140 519 bp mRNA EST 20-APR-1997
LOCUS EST28097 Cerebellum II Homo sapiens cDNA 5' end similar to tyrosine
DEFINITION kinase, receptor (GB:L1315), mRNA sequence.
ACCESSION AA325140
NID 91977395
VERSION AA325140.1 GI:1977395
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 519)
Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Georghagen,N.S.,
Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr.,
Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
Moreno-Palancas,R.F., McDonald,D.A., Nguyen,D.T., Pelligrino,S.M.,
Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H.,
Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
Fraser,C.M. and Venter,J.C.
Initial assessment of human gene diversity and expression patterns
based upon 83 million nucleotides of cDNA sequence
Nature 377 (5547 Suppl), 3-174 (1995)
96026280
On Nov 29, 1993 this sequence version replaced gi:430338.
Other_ESTs: TH017728
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlavet@tigr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (http://www.tigr.org/tdb/hgi/hgi.html)
Seq primer: M13 Reverse.
Location/Qualifiers
1. 519
/organism="Homo sapiens"
/db.xref="ATCC (inhost):125647"
/db.xref="taxon:9606"
/clone_lib="Cerebellum II"
/tissue_type="cerebellum"
/dev_stage="adult"
/note="Organ: brain; Vector: pBluescript SK-; Site:1:
EcoRI; Site:2: XhoI"
117 a 145 c 144 g 108 t 5 others

```

```

Query Match 13.0%; Score 471.6; DB 31; Length 519;
Best Local Similarity 97.7%; Pred. No. 4.4e-84;
Matches 507; Conservative 0; Mismatches 9; Indels 3; Gaps 3;

QY 1965 GGTGCAGAGCCCTCAAGATCTGTGCTGATTTTCCCTTAATGTGCTGAAGGACA 2024
Db 1 GGTGCAGAGCCCTCAAGATCTGTGCTGATTTTCCCTTAATGTGCTGAAGGACA 60
QY 2025 CCCTTTGCTGGTGTGCTCAAGATCTTTAGCGGCAGATGCCACCAAGATGCCAGGAATNA 2084
Db 61 CCCTTTGCTGGTGTGCTCAAGATCTTTAGCGGCAGATGCCACCAAGATGCCAGGAATNA 120
QY 2085 TTTCCCTGAAAGAGGTGAAGATCATGTGCGAGGCTCAAGAGCCCAAAACATCATTCGGCTGCT 2144
Db 121 TTTCCCTGAAAGAGGTGAAGATCATGTGCGAGGCTCAAGAGCCCAAAACATCATTCGGCTGCT 180
QY 2145 GGGCGTGTGTGTCAGGACGACCCCTCTGCTGATGATTACTGACTACATGGAGAACGGCA 2204
Db 181 GGGCGTGTGTGTCAGGACGACCCCTCTGCTGATGATTACTGACTACATGGAGAACGGCA 240
QY 2205 CCTCAACAGTTCTCTCAGTCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2264
Db 241 CCTCAACAGTTCTCTCAGTCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
QY 2265 GGACGGCGAGGCTCGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2323
Db 301 GGACGGCGAGGCTCGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
QY 2324 CCAGATCGGCTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2383
Db 361 CCAGATCGGCTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
QY 2384 CCACGGGAGGCTCGGCTAGTT-GGGGAAATTTTCAACATCAAAATCGACAGCTTTGGCATG 2442
Db 421 NCAAGCGGAGGCTCGGCTAGTTGGGGGAAATTTTCAACATCAAAATCGACAGCTTTGGCATG 480
QY 2443 AGCGGAGGCTCTA-TGCTGGGAGCTATTACCGTGTGCA 2480
Db 481 AGCGGAGGCTCTA-TGCTGGGAGCTATTACCGTGTGCA 519

RESULT 7
AI204173 474 bp mRNA EST 28-OCT-1998
LOCUS qd70e12.xl Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1734862
DEFINITION 3', mRNA sequence.
ACCESSION AI204173
NID 93756779
VERSION AI204173.1 GI:3756779
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 474)
NCI-CCAP http://www.ncbi.nlm.nih.gov/nciccap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
On Jan 17, 1998 this sequence version replaced gi:2044289.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 675 Std Error: 0.00

```

Seq primer: -40UP from Gibco
High quality sequence stop: 395.
Location/Qualifiers
1. .474

FEATURES

/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="2"
/clone="IMAGE:1734862"
/clone_lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"

/note="vector: p77T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer [5', and cloned into the Not I and Eco RI sites of the modified p77T3 vector. Library went through one round of normalization to Cot5, and was constructed by Bento Soares and M. Fatima Bonaldo." 130 a 116 c 121 g 107 t

BASE COUNT

ORIGIN

Query Match 12.9%; Score 470.8; DB 43; Length 474;
Best Local Similarity 99.6%; Pred. No. 6.1e-84;
Matches 472; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3123 CTGCGCCCCAGCTTCTCTCCCTGTGCACACCTGGACCCACCTGGCTGAGATCTGGG 3182
Db 474 CTGCGCCCCAGCTTCTCTCCCTGTGCACACCTGGACCCACCTGGCTGAGATCTGGG 415

QY 3183 GGTGAGGAGCAAGAGGAGGAGAAATGTTCTCTGTGCTGCTGCTGCTGCTGCT 3242
Db 414 GTGAGGAGGAGCAAGAGGAGGAGAAATGTTCTCTGTGCTGCTGCTGCTGCTGCT 355

QY 3243 CAGCTTGGGCTTCTCTCTCCATCCTCACTGAACACTGGACCTGGGGGTAGCCCGGCC 3302
Db 354 CAGCTTGGGCTTCTCTCTCCATCCTCACTGAACACTGGACCTGGGGGTAGCCCGGCC 295

QY 3303 CAGCCCTCAGTACCCGCCACTTCCACCTGCAGTCTTGTAGTAGAATCTCTAAGCCT 3362
Db 294 CAGCCCTCAGTACCCGCCACTTCCACCTGCAGTCTTGTAGTAGAATCTCTAAGCCT 235

QY 3363 ATACGTTCTGTGGAGTAATATTTGGGATTTGGGGAAGAGGAGGAGCAACGCCCATAGC 3422
Db 234 ATACGTTCTGTGGAGTAATATTTGGGATTTGGGGAAGAGGAGGAGCAACGCCCATAGC 175

QY 3423 CTTGGGGTTGGACATCTCTAGTGTAGTCCACATTTATTTTCTATATATCACTTGGGT 3482
Db 174 CTTGGGGTTGGACATCTCTAGTGTAGTCCACATTTATTTTCTATATATCACTTGGGT 115

QY 3483 TTGTACATTTTGGGGGAGAGACACAGATTTTACACTAATATATGACCTAGCTTGAG 3542
Db 114 TTGTACATTTTGGGGGAGAGACACAGATTTTACACTAATATATGACCTAGCTTGAG 55

QY 3543 GCAATTTTAAATCCCTGCACTAGGAGGTAATAATAAGGTTGAGTTTTCACCA 3596
Db 54 GCAATTTTAAATCCCTGCACTAGGAGGTAATAATAAGGTTGAGTTTTCACCA 1

RESULT 8

AI365945

LOCUS

DEFINITION

AI365945 464 bp mRNA EST 07-JAN-1999

IMAGE:1953198.3' similar to SW-EDDL_HUMAN Q08345 EPIHELIAL

DISCOIDIN DOMAIN RECEPTOR 1 PRECURSOR ; mRNA sequence.

AI365945

94125634

AI365945.1 GI:4125634

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Homidae; Homo.

1 (bases 1 to 464)

Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,

Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M.,

Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F.,

Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.

WashU-NCI human EST Project

Unpublished (1997)

On Jan 17, 1998 this sequence version replaced gi:2043665.

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@wustl.edu

This clone is available royalty-free through LNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Possible reversed clone: similarity on wrong strand

Possible reversed clone: polyt not found

Seq primer: -40UP from Gibco.

Location/Qualifiers

1. .464

/organism="Homo sapiens"

/db_xref="taxon:9606"

/map="20"

/clone="IMAGE:1953198"

/clone_lib="Schiller meningioma"

/sex="female"

/tissue_type="meningioma"

/dev_stage="72 years"

/lab_host="SOLR"

/note="Organ: brain; Vector: pBluescript SK- (Stratagene);

Site_1: EcoRI; Site_2: XhoI; Double-stranded cDNA was

prepared from human meningioma using primer

5'-GAGAGAGAGAGAGAGAGAACTGCTGAGT(18)-3'. An EcoRI

adaptor was used on the 5' end of the cDNA as follows:

5'-AATTCGGCAGCAG-3'. The library was size-selected and

went through one round of amplification. Average insert

size is 1.7 kb, with a range from 0.4-12 kb. Tumor

identification by consensus pathology. This library was

constructed by Dr. Martin Schiller (Johns Hopkins

University)."

BASE COUNT 85 a 134 c 155 g 90 t

ORIGIN

Query Match 12.6%; Score 457.6; DB 45; Length 464;

Best Local Similarity 99.1%; Pred. No. 2.5e-81;

Matches 460; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2225 CCCACCACTGAGGAGCAAGCGAGCGGCGCCCTGGGACGCGGCTGCGCAGG 2284

Db 1 CCCACCACTGAGGAGCAAGCGAGCGGCGCCCTGGGACGCGGCTGCGCAGG 60

QY 2285 GGCCCACTCAGTACCCCAATGCTGCTGCATGTGGCAGCCAGATGCGCTCCGCGATGC 2344

Db 61 GGCCCACTCAGTACCCCAATGCTGCTGCATGTGGCAGCCAGATGCGCTCCGCGATGC 120

QY 2345 GCTATCTGGCCACACTCAACTTTGTACATCGGACCTGGCCACCGAACTGCTAGTTG 2404

Db 121 GCTATCTGGCCACACTCAACTTTGTACATCGGACCTGGCCACCGAACTGCTAGTTG 180

QY 2405 GGGAAATTTTCCACCATCAAAATCGCAGACTTTGGCATGAGCCGGAACCTCTATCTGGGG 2464

Db 181 GGGAAATTTTCCACCATCAAAATCGCAGACTTTGGCATGAGCCGGAACCTCTATCTGGGG 240

QY 2465 ACTATTACCGTGTCCAGGGCGGCGGCGAGTGTGCCATCCCTGGATGCCCTGGAGTGCA 2524

Db 241 ACTATTACCGTGTCCAGGGCGGCGGCGAGTGTGCCATCCCTGGATGCCCTGGAGTGCA 300

QY	2525	TCCTCATGGGGAATTCACGACTGCGAGTGTGGGCGCTTTGGTGTGACCGCTGTGGG	2584			
Db	301	TCCTCATGGGGAATTCACGACTGCGAGTGTGGGCGCTTTGGTGTGACCGCTGTGGG	360			
QY	2585	AGTGCTGATGCTCTGTAGGGCCACGCCCTTTGGCAGCTCACCGACGACGAGTCATCG	2644			
Db	361	AGTGCTGATGCTCTGTAGGGCCACGCCCTTTGGCAGCTCACCGACGACGAGTCATCG	420			
QY	2645	AGAACCGGGGGAGTCTTCCTCGGGACACGAGCGCGGACGAGGTAC	2688			
Db	421	AGAACCGGGGGAGTCTTCCTCGGGACACGAGCGCGGACGAGGTAC	464			
RESULT 9						
LOCUS	AA116659	532 bp	EST 13-FEB-1997			
DEFINITION	mq23f05.r1 Barstead MPLRB1 Mus musculus cDNA clone IMAGE:579585 5', similar to TR:G1161063 G1161063 RECEPTOR KINASE. ;, mRNA sequence.					
ACCESSION	AA116659					
NID	g1671670					
KEYWORDS	AA116659.1	GI:1671670				
SOURCE	EST.					
ORGANISM	Mus musculus					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.					
AUTHORS	1 (bases 1 to 532)					
	Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.					
TITLE	The WashU-HMI Mouse EST Project					
JOURNAL	Unpublished (1996)					
COMMENT	On Sep 12, 1996 this sequence version replaced gi:1402199.					
Contact: Marra M/Mouse EST Project						
WashU-HMI Mouse EST Project						
Washington University School of Medicine						
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108						
Tel: 314 286 1800						
Fax: 314 286 1810						
Email: mouseest@watson.wustl.edu						
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.						
MGI:354233						
Possible reversed clone: similarity on wrong strand						
Seq primer: -28m13 rev2 Er from Amersham						
High quality sequence stop: 404.						
FEATURES	Location/Qualifiers					
	1..532					
	/organism="Mus musculus"					
	/strain="BALB/c"					
	/db_xref="taxon:10090"					
	/map="815B06; 5; 5q14.1-5q15"					
	/clone="IMAGE:579585"					
	/clone.lib="Barstead MPLRB1"					
	/sex="mixed"					
	/tissue_type="Kidney"					
source	/dev_stage="6 weeks"					
	/lab_host="DH10B"					
	/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: ECORI; Site_2: NotI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAAGTGGGAGCGCGCCCTTTTGTGTGACCGCTGTGGG 3']; double-stranded cDNA was ligated to Eco RI adaptors [CATGCGATTGGGATCC], digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library constructed by Bob Barstead."					
	89 a	209 c	134 g 100 t			
	BASE COUNT					
	ORIGIN					

QY 3509 AGATTTTACACAAATATATATAGGACCTAGCTTGTAGGCAATTTTAAATCCCTGCACCTAGGCA 3568
 |||||
 Db 94 AGATTTTACACAAATATATATAGGACCTAGCTTGTAGGCAATTTTAAATCCCTGCACCTAGGCA 35
 |||||
 QY 3569 GGTAATAATAAGGTTGAGTTTCCACAAAAA 3602
 |||||
 Db 34 GGTAATAATAAGGTTGAGTTTCCAAAAA 1
 |||||

RESULT 12
 AA944928/c 542 bp mRNA EST 16-JUN-1998
 LOCUS EST200427 Normalized rat embryo, Bento Soares Rattus sp. cDNA clone
 DEFINITION REMA085 3' end, mRNA sequence.
 ACCESSION AA944928
 NID 93104844
 VERSION AA944928.1 GI:3104844
 KEYWORDS EST.
 SOURCE Rattus sp.
 ORGANISM Rattus sp.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 REFERENCE 1 (bases 1 to 542)
 AUTHORS Lee,N.H., Glodek,A., Chandra,I., Mason,T.M., Quackenbush,J.,
 Kerlavage,A.R. and Adams,M.D.
 TITLE Rat Genome Project: Generation of a Rat EST (REST) Catalog & Rat
 Gene Index
 JOURNAL Unpublished (1998)
 COMMENT On Jan 17, 1998 this sequence version replaced gi:1900074.

Contact: Lee, NH
 ATCC
 The Institute for Genomic Research
 9712, Medical Center Drive, Rockville, MD 20850, USA
 Tel: (301)-838-3529
 Fax: (301)-838-0208
 Email: nhlee@igir.org
 Seq primer: M13-21.

FEATURES
 Source
 1..542
 /organism="Rattus sp."
 /db_xref="ATCC (inhost):2014392"
 /db_xref="taxon:10118"
 /map="x"
 /clone="REMA085"
 /clone_lib="Normalized rat embryo, Bento Soares"
 /dev_stage="embryo 8, 12, 18 dpc"
 /note="Vector: pT73Pac; Site_1: EcoRI; Site_2: NotI"
 BASE COUNT 110 a 169 c 151 g 111 t 1 others
 ORIGIN

Query Match 11.9%; Score 432.2; DB 40; Length 542;
 Best Local Similarity 87.3%; Pred. No. 2.9e-76;
 Matches 473; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 2221 AGTGCCACAGCTGAGGACAGGACGCGAGGGGGCCCTGGGGAGCGGCGAGCTGCG 2280
 |||||
 Db 542 AGTGCCACAGCTGAGGACAGGACGCTACTAGGCGCTCTCTGGNGACAGAGCTCTGAC 483
 |||||
 QY 2281 CAGGGGCCCCACCATGACTACCATAGCTGCTGCGATGTGGCAGGCCAGATCGCTCCGGC 2340
 |||||
 Db 482 CAGGGGCCCCACCATGACTACCATAGCTGCTGTTACAGCTGGGGGCCAGATCGCTCTGCG 423
 |||||
 QY 2341 ATGCGGTATCTGGCCACACTCACTTTGTACATCGGACCTGGCCACCGGACCTGCCTA 2400
 |||||
 Db 422 ATGCGGTATCTGGCCACCTTAACTTTGTCATCGGACCTGGCCACCGGAACTGCTTG 363
 |||||
 QY 2401 GTTGGGAAAATTTACCATCAAAATCGCAGACTTTGGCATGAGCGGAACTCTATGCT 2460
 |||||
 Db 362 GTTGGGAAAATTTACCATCAAAATCGCGACTTTGGCATGAGCGGAACTCTACGCT 303
 |||||

QY 2461 GGGGACTATTACCGTGTGCGAGGGCCGCGAGTCTGCCCATCCGCTGATGGCCTGGAG 2520
 |||||
 Db 302 GGGGATTATTACCGTGTACAGGGCCGCGCAGTGTGCTGCCATCAGGTGATGGCCTGGAG 243
 |||||
 QY 2521 TGCATCTCATGGGGAAGTTTACAGCTGCGAGTGACCTGTGGGCTTTGGTGTGACCTG 2580
 |||||
 Db 242 TGTATCTCATGGGGAAGTTTACAAACAGCCAGTGTGTGGGCTTCGGAGTGACCTG 183
 |||||
 QY 2581 TGGGAGTGTGCTGTGTAGGGCCCGAGCCCTTTGGGAGCTCACGAGCAGGAGTC 2640
 |||||
 Db 182 TGGGAGTGTGCTGTATGCTGTGAGGTTCAGCCCTTCGGGCACTCACAGCAGGAGTC 123
 |||||
 QY 2641 ATCAGAACCGGGGAGTGTCTTCCGGGACAGGCGCGAGGTGTACCTGTCCCGCCG 2700
 |||||
 Db 122 ATCAGAACCGGGGAGTGTCTTCCAGGAGTGTCTTCCAGGAGTGTATCTATCCAGGCA 63
 |||||
 QY 2701 CCTGCTTCCCGAGGAGGCTTATGAGCTGATGCTTGGTGTGGAGCGGAGTGTGAG 2760
 |||||
 Db 62 CCGGCTTCCCGAGGAGTGTGATGAGCTGATGCTCGGTGTGGAGCGGAGCGGAG 3
 |||||

QY 2761 CA 2762
 ||
 Db 2 CA 1
 ||

RESULT 13
 AA430320 429 bp mRNA EST 16-OCT-1997
 LOCUS zW07fill.r1 Soares_NhMHPu_S1 Homo sapiens cDNA clone IMAGE:768621
 DEFINITION 5', mRNA sequence.
 ACCESSION AA430320
 NID 92110911
 VERSION AA430320.1 GI:2110911
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 429)
 AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
 Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J.,
 Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
 White,Y., Wylie,T., Waterston,R. and Wilson,R.
 TITLE WashU-Merck EST Project 1997
 JOURNAL Unpublished (1997)
 COMMENT On Sep 12, 1996 this sequence version replaced gi:1392778.

Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: -28ml3 rev2 Et from Amersham
 High quality sequence stop: 411.
 Location/Qualifiers
 1..429
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:768621"
 /clone_lib="Soares_NhMHPu_S1"
 /tissue_type="Pooled human melanocyte, fetal heart, and
 pregnant uterus"
 /lab_host="DH10B"

FEATURES
 source

/note="Organ: mixed (see below); Vector: pT73D-Pac
 (Pharmacia) with a modified polylinker; Site_1: Not I;
 Site_2: Eco RI; Equal amounts of plasmid DNA from three
 normalized libraries (melanocyte 2NBHM, pregnant uterus
 NDHPU, and fetal heart NDH19W) were mixed, and ss circles
 were made in vitro. Following HAP purification, this DNA
 was used as tracer in a subtractive hybridization

DNA Sequencing by: Washington University Genome Sequencing Center

EST.

C04008.1 GI:1467259

SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 465)
AUTHORS Tanaka,T., Ogiwara,A., Uchiyama,I., Takagi,T., Yazaki,Y. and Nakamura,Y.
TITLE Construction of a Normalized Directionally Cloned cDNA Library from Adult Heart and Analysis of 3040 Clones by Partial Sequencing
JOURNAL Genomics 35, 231-235 (1996)
COMMENT On Nov 22, 1995 this sequence version replaced gi:1071057.
Contact: Yusuke Nakamura
Institute of Medical Science
University of Tokyo
4-6-1, Shirokanedai, Minato-ku, Tokyo 108, Japan
Tel: 81-3-5449-5372
Fax: 81-3-5449-5433
Email: yusuke@ims.u-tokyo.ac.jp.
Location/Qualifiers
1. .465
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="3NHC2547"
/clone_lib="Human heart cDNA (Y Nakamura)"
/dev_stage="adult"
/note="Organ: heart; normalized directionally cloned cDNA from adult heart"
BASE COUNT 88 a 146 c 111 g 114 t 6 others
ORIGIN

Query Match 11.4%; Score 416; DB 27; Length 465;
Best Local Similarity 98.4%; Pred. No. 4.5e-73;
Matches 427; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

QY 2959 AGAAGCCCTGTGCGCCACCCAGCTGGTCTGTGGATGGGATCCTCTCCACCCCTCCTCTA 3018
|||||
Db 32 AGAAGCCCTGTGCGCCACCCAGCTGGTCTGTGGATGGGATCCTCTCCACCCCTCCTCTA 91
|||||

QY 3019 GCCATCCCTTGGGAGGTTGGGAGAAATATAGGATAGACACTGGACATGCCCATTTGG 3078
|||||
Db 92 GCCATCCCTTGGGAGGTTGGGAGAAATATAGGATAGACACTGGACATGCCCATTTGG 151
|||||

QY 3079 AGCACCCTGGGCCCTACTGGACACACTGATTCCTGGAGAGTGGCTGCGCCCCCAGCTTC 3138
|||||
Db 152 AGCACCCTGGGCCCTACTGGACACACTGATTCCTGGAGAGTGGCTGCGCCCCCAGCTTC 211
|||||

QY 3139 TCCTCTCCCTGTACACACTGGACCCACCTGGCTGAGATCTGGGGGTGAGGAGACAAGA 3198
|||||
Db 212 TCCTCTCCCTGTACACACTGGACCCACCTGGCTGAGATCTGGGGGTGAGGAGACAAGA 271
|||||

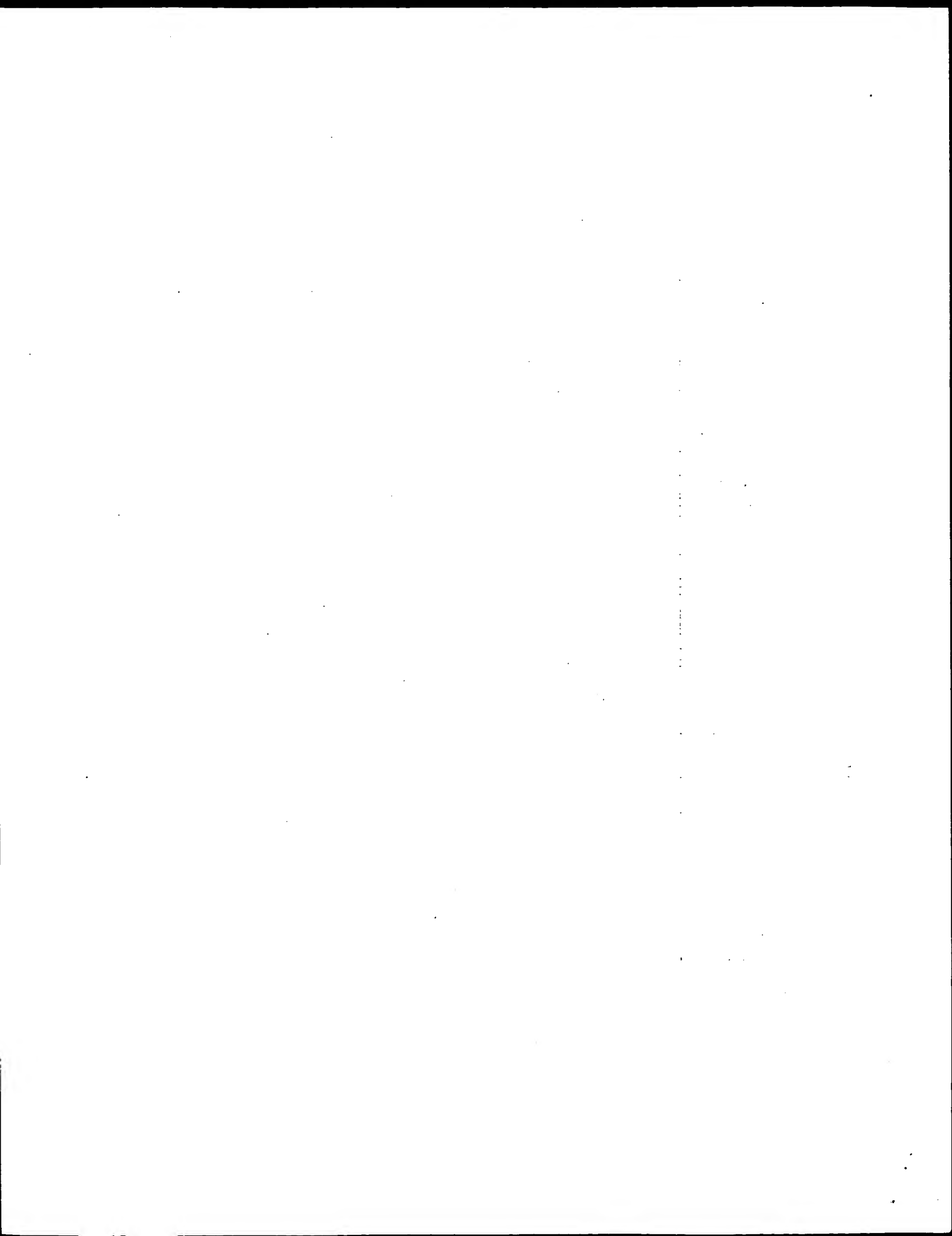
QY 3199 AGGAGAGGAAATGTTTCCTTGTGCTGCTCTGTACTTGTCTCAGCTTGGGCTTCTTC 3258
|||||
Db 272 AGGAGAGGAAATGTTTCCTTGTGCTGCTCTGTACTTGTCTCAGCTTGGGCTT-TTC 330
|||||

QY 3259 CTCCTCCATCACCCTGAACACTGGACCTGGGGGTAGCCCCCAGCCCTCAGTCACCC 3318
|||||
Db 331 CTCCTNCATCACCCTGAACACTGGACCTGGGGGTAGCCCCCAGCCCTCAGTCACCC 390
|||||

QY 3319 CCACCTTCCACCTGGAGTCTTGTAGTAGAATCTCTAAGCCTATACGTTTCTGTGGAG 3378
|||||
Db 391 CNACTTCCACCTGGAGTCTTGTAGTAGAATCTCTAAGCCTATACGTTTCTTNTGGAG 450
|||||

QY 3379 TAAATATGGGATT 3392
|||||
Db 451 TAAATATGGGATT 464
|||||

Search completed: November 4, 1999, 07:05:18
Job time: 4934 sec



Result No.	Query			Description	
	Score	Match	Length	ID	Description
1	3637	100.0	3637	2	US-08-4445-640-3
2	3451	94.9	3962	2	US-08-336-343A-1
3	1137	32.3	1197	2	US-08-445-640-7
4	669.4	18.4	3120	3	US-08-456-647B-19
5	669.4	18.4	3120	3	US-08-237-401A-19
6	666.8	18.3	3157	2	US-08-336-343A-3
7	666.8	18.3	3157	2	US-08-336-343A-5
8	195.2	5.4	2820	3	US-08-286-305A-4
9	195.2	5.4	2820	4	US-08-441-104A-4
10	195.2	5.4	2820	4	US-08-440-816A-4
11	193.6	5.3	2301	2	US-08-306-691B-23
12	193.6	5.3	2301	5	PCR-US93-06251-78
13	190	5.2	3707	1	US-08-271-454-1
14	190	5.2	3060	3	US-08-286-305A-6
15	190	5.2	3194	3	US-08-359-705B-1
16	190	5.2	3194	4	US-08-286-846A-1
17	190	5.2	3060	4	US-08-441-104A-6
18	190	5.2	3194	4	US-08-457-880A-1
19	190	5.2	3060	4	US-08-440-816A-6
20	190	5.2	3707	5	PCR-US95-08180-1
21	185.2	5.1	2940	3	US-08-286-305A-8
22	185.2	5.1	2940	4	US-08-441-104A-8
23	185.2	5.1	2940	4	US-08-440-816A-8
24	182	5.0	2526	1	US-07-912-952-1
25	172.4	4.7	2463	1	US-08-339-578-1
26	157.4	4.3	3398	5	PCR-US95-08493-12
27	156.4	4.3	4092	3	US-08-469-537A-106
28	144.4	4.0	2208	5	PCR-US95-08493-1
29	144.4	4.0	2580	5	PCR-US95-08493-18
30	144.4	4.0	2604	5	PCR-US95-08493-20
31	144.2	4.0	2376	1	US-07-912-952-3
32	140.6	3.9	2869	1	US-08-374-834-2
33	140.6	3.9	2869	3	US-08-644-271-2
34	135.2	3.7	4149	4	US-08-737-715-1
35	133.2	3.7	2715	3	US-08-359-705B-5
36	133.2	3.7	2715	4	US-08-286-846A-5
37	133.2	3.7	2715	4	US-08-457-880A-5

Db 61 AGGCCCCAGGGATCAGAGCTATGGACACAGAGCCCTGTCATCTTACTGTGCTCTC 120
Qy 121 TTGTTGCAAGTGTGAGATGCTGACATGAAGGACATTTTGTATCCTGCCAAGTGCCGCTAT 180
Db 121 TTGTTGCAAGTGTGAGATGCTGACATGAAGGACATTTTGTATCCTGCCAAGTGCCGCTAT 180
Qy 181 GCGCTGGGCTATGACAGGACCGGACCATCCACAGACAGTGACATCTCTCTCCAGCTCCTGG 240
Db 181 GCGCTGGGCTATGACAGGACCGGACCATCCACAGACAGTGACATCTCTCTCCAGCTCCTGG 240
Qy 241 TCAGATTCCACTGCGGCCCCACAGCAGGTGGAGAGCAGTGACGGGATGGGCTGG 300
Db 241 TCAGATTCCACTGCGGCCCCACAGCAGGTGGAGAGCAGTGACGGGATGGGCTGG 300
Qy 301 TGCCCCCAGGGTGGGTGTTTCCAGGAGGAGGAGTACTTTCAGGTGATCTACAAAGA 360
Db 301 TGCCCCCAGGGTGGGTGTTTCCAGGAGGAGGAGTACTTTCAGGTGATCTACAAAGA 360
Qy 361 CTGACCTGGTGGCTGTGGGACCCAGGACCGCATGCCGGGGCTGGGCAAGAG 420
Db 361 CTGACCTGGTGGCTGTGGGACCCAGGACCGCATGCCGGGGCTGGGCAAGAG 420
Qy 421 TTCTCCGGAGCTACCGGCTGGTTACTCCGGGATGGTCGCCGCTGGATGGGCTGGAAG 480
Db 421 TTCTCCGGAGCTACCGGCTGGTTACTCCGGGATGGTCGCCGCTGGATGGGCTGGAAG 480
Qy 481 GACCGCTGGGTGAGAGGTGATCTCAGGCAATGAGGACCTCAGGAGGTGGTGTGAAG 540
Db 481 GACCGCTGGGTGAGAGGTGATCTCAGGCAATGAGGACCTCAGGAGGTGGTGTGAAG 540
Qy 541 GACCTTGGGCCCCCATGGTTGCCCGACTGTTTCGCTTACCCCGGGCTGACCGGGTC 600
Db 541 GACCTTGGGCCCCCATGGTTGCCCGACTGTTTCGCTTACCCCGGGCTGACCGGGTC 600
Qy 601 ATGAGCTGTCTGCGGGGTAGAGCTCTATGGCTGCTCTGGAGGATGGATCTCTGTCT 560
Db 601 ATGAGCTGTCTGCGGGGTAGAGCTCTATGGCTGCTCTGGAGGATGGATCTCTGTCT 560
Qy 661 TACACCGCCCTGTGGGGCAGACAATGATTTATCTGAGGCCGTGACCTCAACGACTCC 720
Db 661 TACACCGCCCTGTGGGGCAGACAATGATTTATCTGAGGCCGTGACCTCAACGACTCC 720
Qy 721 ACCTATGACGACATACCGTGGGCGGACTGCAGTATGGGGTCTGGGCCAGCTGGCAGAT 780
Db 721 ACCTATGACGACATACCGTGGGCGGACTGCAGTATGGGGTCTGGGCCAGCTGGCAGAT 780
Qy 781 GGTGTGGGGCTGAGTACTTTAGGAAGAGTCAGGAGCTGGGGTCTGGGCCAGGCTAT 840
Db 781 GGTGTGGGGCTGAGTACTTTAGGAAGAGTCAGGAGCTGGGGTCTGGGCCAGGCTAT 840
Qy 841 GACTATGTGGGATGGAGCAACACAGCTTCTCCAGTGGCTATGTGGAGATGGAGTTTGA 900
Db 841 GACTATGTGGGATGGAGCAACACAGCTTCTCCAGTGGCTATGTGGAGATGGAGTTTGA 900
Qy 901 TTTGACCGGCTGAGGGCTTCCAGGCTATGAGTCCAGTCCAGTCAACACATGACACGCTG 960
Db 901 TTTGACCGGCTGAGGGCTTCCAGGCTATGAGTCCAGTCCAGTCAACACATGACACGCTG 960
Qy 961 GGAGCCGCTGCTGCGGGGTGGAAATGTCGCTTCCGGGCTGGGCTGGCCATGGGCTGG 1020
Db 961 GGAGCCGCTGCTGCGGGGTGGAAATGTCGCTTCCGGGCTGGGCTGGCCATGGGCTGG 1020
Qy 1021 GAGGGGAGCCCATGCGCCACAACCTAGGGGCAACCTGGGGGACCCACAGAGCCGGGCT 1080
Db 1021 GAGGGGAGCCCATGCGCCACAACCTAGGGGCAACCTGGGGGACCCACAGAGCCGGGCT 1080
Qy 1081 GTCTAGTGGCCCTTGGCGCGGTGCTGCTTCTGTCAGTGGCGCTTCTCTTTTGG 1140
Db 1081 GTCTAGTGGCCCTTGGCGCGGTGCTGCTTCTGTCAGTGGCGCTTCTCTTTTGG 1140
Qy 1141 GGGCCCTGGTTACTCTTACGGGAATCTCTCTATCTCTGATGTGGTGAACAATTCCTCT 1200
Db 1141 GGGCCCTGGTTACTCTTACGGGAATCTCTCTATCTCTGATGTGGTGAACAATTCCTCT

Db 1141 GGGCCCTGGTTACTCTTACGGGAATCTCTCTATCTCTGATGTGGTGAACAATTCCTCT 1200
Qy 1201 CCGCACTGGAGGACACCTTCCCGCCAGCCCCCTGTGTGGCGCTGCCACCTCCCAACC 1260
Db 1201 CCGCACTGGAGGACACCTTCCCGCCAGCCCCCTGTGTGGCGCTGCCACCTCCCAACC 1260
Qy 1261 AACTTACAGCTTGGAGCTGGAGCCAGAGGCCAGAGCCCGTGGCCAAAGCCCGAGGGG 1320
Db 1261 AACTTACAGCTTGGAGCTGGAGCCAGAGGCCAGAGCCCGTGGCCAAAGCCCGAGGGG 1320
Qy 1321 AGCCGACCGCATCTCATCGGCTGCTGTGGCCATCATCTCTCTCTCTCTCTCTCTCT 1380
Db 1321 AGCCGACCGCATCTCATCGGCTGCTGTGGCCATCATCTCTCTCTCTCTCTCTCTCT 1380
Qy 1381 ATTGCGCTCATCTCTGCGGCTGCACTGGCGAGGCTCTCAGCAAGGCTGAACGAGG 1440
Db 1381 ATTGCGCTCATCTCTGCGGCTGCACTGGCGAGGCTCTCAGCAAGGCTGAACGAGG 1440
Qy 1441 GTGTTGGAAGAGGAGCTGACGGTTTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1500
Db 1441 GTGTTGGAAGAGGAGCTGACGGTTTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1500
Qy 1501 AACCGCCAGCTCTAGAGAGCCACCCCTACAGAGCCCGGCTCTCTGGGAATCCG 1560
Db 1501 AACCGCCAGCTCTAGAGAGCCACCCCTACAGAGCCCGGCTCTCTGGGAATCCG 1560
Qy 1561 CCCCACTCCGCTCTCTGTCCTCCCAATGGCTCTGCTCTCTCTCTCTCTCTCTCTCTCT 1620
Db 1561 CCCCACTCCGCTCTCTGTCCTCCCAATGGCTCTGCTCTCTCTCTCTCTCTCTCTCTCT 1620
Qy 1621 CGCTCTCTCTGCGGCTTACGGCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1680
Db 1621 CGCTCTCTCTGCGGCTTACGGCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1680
Qy 1681 TGGGCAAAACCCACACAGCCAGGCTTACAGTGGGACTATATGGAGCTGAGAAGCCA 1740
Db 1681 TGGGCAAAACCCACACAGCCAGGCTTACAGTGGGACTATATGGAGCTGAGAAGCCA 1740
Qy 1741 GCGCCCCGCTTCTGCCCCACCTCCCGCAGAACAGCTGCCCCATATGCGGAGGCTGAC 1800
Db 1741 GCGCCCCGCTTCTGCCCCACCTCCCGCAGAACAGCTGCCCCATATGCGGAGGCTGAC 1800
Qy 1801 ATTGTTACCTTGCAGGGGCTCACCAGGAGAACACCTATGCTGTGCTGCACTGCCGCCA 1860
Db 1801 ATTGTTACCTTGCAGGGGCTCACCAGGAGAACACCTATGCTGTGCTGCACTGCCGCCA 1860
Qy 1861 GGGCAGCTCGGGATGGCCCCCAGAGTGGATTTCCCTCGATCTCGACTCCGCTTCAAG 1920
Db 1861 GGGCAGCTCGGGATGGCCCCCAGAGTGGATTTCCCTCGATCTCGACTCCGCTTCAAG 1920
Qy 1921 GAGAGCTTGGCGAGGGCCAGTTTGGGAGGTGACACCTGTGTGAGGTGCGAGCCCTCAA 1980
Db 1921 GAGAGCTTGGCGAGGGCCAGTTTGGGAGGTGACACCTGTGTGAGGTGCGAGCCCTCAA 1980
Qy 1981 GATCTGTGCTGTGATTTCCCTTAAATGTGGTAAAGGACACCTTTTGTGGTGTAGT 2040
Db 1981 GATCTGTGCTGTGATTTCCCTTAAATGTGGTAAAGGACACCTTTTGTGGTGTAGT 2040
Qy 2041 GTCAGATCTTACGGCCAGATGCCACCAAGATGCCAGGAATGATTTCTTAAAGAGGTG 2100
Db 2041 GTCAGATCTTACGGCCAGATGCCACCAAGATGCCAGGAATGATTTCTTAAAGAGGTG 2100
Qy 2101 AAGATCATGTGAGGCTCAAGGACCCCAACATCATTTCCGCTGCTGGGCTGTGTGTGAG 2160
Db 2101 AAGATCATGTGAGGCTCAAGGACCCCAACATCATTTCCGCTGCTGGGCTGTGTGTGAG 2160
Qy 2161 GACGACCCCTTCTGATGATTTACTGACTACATGAGAACGCGACCTCAACAGCTTCTCT 2220
Db 2161 GACGACCCCTTCTGATGATTTACTGACTACATGAGAACGCGACCTCAACAGCTTCTCT 2220
Qy 2221 AGTGGCCACAGCTGGAGGCAAGCAGCCAGGGGCCCCCTGGGAGGCGGCTGGG 2280
Db 2221 AGTGGCCACAGCTGGAGGCAAGCAGCCAGGGGCCCCCTGGGAGGCGGCTGGG 2280

QY 2281 CAGGGGCCACCATCAGCTACCAATGCTGCTGATGTGGCAGCCAGATCGCCTCCGGC 2340
DB 2281 CAGGGGCCACCATCAGCTACCAATGCTGCTGATGTGGCAGCCAGATCGCCTCCGGC 2340
QY 2341 ATGGCTATCTGGCCACACCTCACTTTGTATCTGGGACCTGGCCAGCGGAACTGCCTA 2400
DB 2341 ATGGCTATCTGGCCACACCTCACTTTGTATCTGGGACCTGGCCAGCGGAACTGCCTA 2400
QY 2401 GTTGGGAAATTTACCATCAAAATCGCAGACTTTGGCATAGCGCGAACTCTATGCT 2460
DB 2401 GTTGGGAAATTTACCATCAAAATCGCAGACTTTGGCATAGCGCGAACTCTATGCT 2460
QY 2461 GGGGACTATTACCGTGTGGAGCGCGGCGAGTCTGCCCATCCGCTGGATGGGAG 2520
DB 2461 GGGGACTATTACCGTGTGGAGCGCGGCGAGTCTGCCCATCCGCTGGATGGGAG 2520
QY 2521 TGCATCTCATGGGAGTTTCAGACTGCGAGTGCAGTGTGGGCTTTGGTGTGACCCCTG 2580
DB 2521 TGCATCTCATGGGAGTTTCAGACTGCGAGTGCAGTGTGGGCTTTGGTGTGACCCCTG 2580
QY 2581 TGGGAGTGTGATGCTCTGTAGGGCCCGCCCTTTGGGAGCTCACCGACGAGCAGTGC 2640
DB 2581 TGGGAGTGTGATGCTCTGTAGGGCCCGCCCTTTGGGAGCTCACCGACGAGCAGTGC 2640
QY 2641 ATCGAGACCGGGGAGTTCTTCGGGACCGAGGCGCGCAGGTGTACCTGTCGCGCGG 2700
DB 2641 ATCGAGACCGGGGAGTTCTTCGGGACCGAGGCGCGCAGGTGTACCTGTCGCGCGG 2700
QY 2701 CTTGCTGCTGCGGAGGCTTATAGCTGATGCTTCTGCTGCTGGGCGGAGTCTGAG 2760
DB 2701 CTTGCTGCTGCGGAGGCTTATAGCTGATGCTTCTGCTGCTGGGCGGAGTCTGAG 2760
QY 2761 CAGCGACCCCTTTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2820
DB 2761 CAGCGACCCCTTTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2820
QY 2821 TGAATCACATCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2880
DB 2821 TGAATCACATCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2880
QY 2881 AAAAAGAGGACACATGCGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2940
DB 2881 AAAAAGAGGACACATGCGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2940
QY 2941 TAGAGGAGTGTGAGCTGACAGAGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3000
DB 2941 TAGAGGAGTGTGAGCTGACAGAGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3000
QY 3001 CTTCTCCACCCCTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3060
DB 3001 CTTCTCCACCCCTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3060
QY 3061 CTTGACATGGCCCATTTGGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3120
DB 3061 CTTGACATGGCCCATTTGGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3120
QY 3121 GGTGCGCCCGCCAGCTTCTG 3180
DB 3121 GGTGCGCCCGCCAGCTTCTG 3180
QY 3181 GGGGTGAGGACAGAGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3240
DB 3181 GGGGTGAGGACAGAGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3240
QY 3241 CTCAGCTTGGGCTTCTG 3300
DB 3241 CTCAGCTTGGGCTTCTG 3300
QY 3301 CCCAGCCCTCAGTACCCCGCCCTG 3360
DB 3301 CCCAGCCCTCAGTACCCCGCCCTG 3360

QY 3361 CTATACGTTTCTGTGGAGTAATATTGGGATTGGGGAGAGAGGAGGAGGAGGAGGAGGAGG 3420
DB 3361 CTATACGTTTCTGTGGAGTAATATTGGGATTGGGGAGAGAGGAGGAGGAGGAGGAGGAG 3420
QY 3421 GCCTTGGGTTGGACATCTCTAGTGTAGTCCACATTTGATTTTCTATATCACTTGGG 3480
DB 3421 GCCTTGGGTTGGACATCTCTAGTGTAGTCCACATTTGATTTTCTATATCACTTGGG 3480
QY 3481 GTTGTACATTTTGGGGGAGAGACACAGATTTTACACTAATATATGGACCTAGCTTG 3540
DB 3481 GTTGTACATTTTGGGGGAGAGACACAGATTTTACACTAATATATGGACCTAGCTTG 3540
QY 3541 AGGCAATTTTAAATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3600
DB 3541 AGGCAATTTTAAATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3600
QY 3601 AAAAAAATTTTAAATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3637
DB 3601 AAAAAAATTTTAAATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3637

RESULT 2

US-08-336-343A-1
; Sequence 1, Application US/08336343A
; Patent No. 5677144
; GENERAL INFORMATION:
; APPLICANT: Ullrich, Axel
; APPLICANT: Alves, Frauke
; TITLE OF INVENTION: CCK-2, A No. 5677144e1 Receptor Tyrosine Kinase
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/336,343A
; FILING DATE: 08-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7683-065
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3962 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 321...3077
; US-08-336-343A-1

Query Match 94.9%; Score 3451; DB 2; Length 3962;
Best Local Similarity: 97.0%; Pred. No. 0;
Matches 3589; Conservative 0; Mismatches 5; Indels 105; Gaps 3;

QY 17 GTTGACTTGAAGAAATGCCAAGAGATGCTGCCCCACCCTTAGCCCGGAGGATCAG 76
DB 256 GTTGACTTGAAGAAATGCCAAGAGATGCTGCCCCACCCTTAGCCCGGAGGATCAG 315
QY 77 GAGCTATGGGACCAAGAGCCCTGTCTATCTTTACTGTGCTGCTCTTGGTGGCAAGTGGAG 136
DB 316 GAGCTATGGGACCAAGAGCCCTGTCTATCTTTACTGTGCTGCTCTTGGTGGCAAGTGGAG 375
QY 137 ATGTGACATCAAGGAGACATTTGATCTGCTGCAAGTGGCTGATGCCCTGGGATGCAGG 196
DB 376 ATGTGACATCAAGGAGACATTTGATCTGCTGCAAGTGGCTGATGCCCTGGGATGCAGG 435
QY 197 ACCGACCATCCAGACAGTACATCTCTGCTCCAGTCTCTGCTGATGCTGATGCTGATGCTGCTG 256
DB 436 ACCGACCATCCAGACAGTACATCTCTGCTCCAGTCTCTGCTGATGCTGATGCTGATGCTGCTG 495
QY 257 CCGGCCACAGAGGTTGGAGAGCAGTACGAGGGGATGGGGCTTGGTGGCCCGCAGGGTGGG 316
DB 496 CCGGCCACAGAGGTTGGAGAGCAGTACGAGGGGATGGGGCTTGGTGGCCCGCAGGGTGGG 555
QY 317 TGTCTTCCCAAGGAGGAGTACTTGCAGGTGGATCTACACGACTGCACCTGGTGGCTC 376
DB 556 TGTCTTCCCAAGGAGGAGTACTTGCAGGTGGATCTACACGACTGCACCTGGTGGCTC 615
QY 377 TGGTGGGACCCAGGACGGATGCGGGGCTTGGCAAGGATTTCTCCCGGAGCTACC 436
DB 616 TGGTGGGACCCAGGACGGATGCGGGGCTTGGCAAGGATTTCTCCCGGAGCTACC 675
QY 437 GCGTCGGTTACTCCCGGATGGTGGCCCTGGATGGGCTGGAAGACCGCTGGGGTTCAGG 496
DB 676 GCGTCGGTTACTCCCGGATGGTGGCCCTGGATGGGCTGGAAGACCGCTGGGGTTCAGG 735
QY 497 AGGTGATCTCAGGCAATGAGGACCTGAGGAGTGGTCTGAAGGACCTTGGGCCCCCA 556
DB 736 AGGTGATCTCAGGCAATGAGGACCTGAGGAGTGGTCTGAAGGACCTTGGGCCCCCA 795
QY 557 TGGTGGCCGACTGCTTCCCTTCTACCCCGGGCTGACCGGCTATGAGCTGTCTGTGTC 616
DB 796 TGGTGGCCGACTGCTTCCCTTCTACCCCGGGCTGACCGGCTATGAGCTGTCTGTGTC 855
QY 617 GGGTAGAGCTCTATGGTGGCTCTGGAGGATGGTCTGCTTACACCGCCCTGTGG 676
DB 856 GGGTAGAGCTCTATGGTGGCTCTGGAGGATGGTCTGCTTACACCGCCCTGTGG 915
QY 677 GGCAGACAATGATTTATCTGAGGCGGTGACTTCAACGAGCTTCCACCTATGACGACATA 736
DB 916 GGCAGACAATGATTTATCTGAGGCGGTGACTTCAACGAGCTTCCACCTATGACGACATA 975
QY 737 CCGTGGCGGACTGCAGTATGGGGTCTGGCCAGCTGGCAGATGGTGTGGGGCTGG 796
DB 976 CCGTGGCGGACTGCAGTATGGGGTCTGGCCAGCTGGCAGATGGTGTGGGGCTGG 1035
QY 797 ATGACTTTAGGAAGTCAAGAGCTGCGGCTCTGCGGAGTATGAGTATGAGGATGGA 856
DB 1036 ATGACTTTAGGAAGTCAAGAGCTGCGGCTCTGCGGAGTATGAGTATGAGGATGGA 1095
QY 857 GCAACACAGCTTCTCCAGTGGCTATGTGGAGATGGATTTGAGTTTACCGGCTGAGGG 916
DB 1096 GCAACACAGCTTCTCCAGTGGCTATGTGGAGATGGATTTGAGTTTACCGGCTGAGGG 1155
QY 917 CCTTCCAGGCTATGAGGTCCACTGTAAACATGCACACGCTGGAGCCGCTGCGCTG 976
DB 1156 CCTTCCAGGCTATGAGGTCCACTGTAAACATGCACACGCTGGAGCCGCTGCGCTG 1215
QY 977 GCGGGTGAATCTGCTTCCGCGTGGCCCTGCCATGGCTGGGAGGGAGCCCATGC 1036
DB 1216 GCGGGTGAATCTGCTTCCGCGTGGCCCTGCCATGGCTGGGAGGGAGCCCATGC 1275
QY 1037 GCCACAACCTAGGGGCAACCTGGGGACCCAGACCGCGGCTGTCTAGTGGCCCTTG 1096
DB 1276 GCCACAACCTAGGGGCAACCTGGGGACCCAGACCGCGGCTGTCTAGTGGCCCTTG 1335
QY 1097 GCGGGCTGTGGCTGCTTCTGCAAGTGGCTTCTCTTTCGGGGGCTTACTCT 1156

DB 1336 GCGGCGGTGTGGCTGCTTCTGCAAGTGGCTTCTCTTTGCGGGGCTTGGTACTCT 1395
QY 1157 TCAGGGAATCTCTCTCTGATGTGGTGAACAAATCTCTCCGGCACTGGGAGCA 1216
DB 1396 TCAGGGAATCTCTCTCTGATGTGGTGAACAAATCTCTCCGGCACTGGGAGCA 1455
QY 1217 CCGTCCCGGACGCGCCCTGGTGGCGGCTTGGCCACCTCCACCAACTTCAGCAGCTGG 1276
DB 1456 CCGTCCCGGACGCGCCCTGGTGGCGGCTTGGCCACCTCCACCAACTTCAGCAGCTGG 1515
QY 1277 AGCTGGAGCCAGAGGCGCAGAGCCGCTGGCAAGCCCGAGGGAGCCCGCCCATCC 1336
DB 1516 AGCTGGAGCCAGAGGCGCAGAGCCGCTGGCAAGCCCGAGGGAGCCCGCCCATCC 1575
QY 1337 TCATGGCTGCTGCTGGCCATCTCTGCTGCTCTCATCTATCTGCTCTCATCTCT 1396
DB 1576 TCATGGCTGCTGCTGGCCATCTCTGCTGCTCTCATCTATCTGCTCTCATCTCT 1635
QY 1397 GCGGCTGCACTGGCGAGGCTCTCAGCAAGGCTGAACGAGGCTGTGGAGAGGAGC 1456
DB 1636 GCGGCTGCACTGGCGAGGCTCTCAGCAAGGCTGAACGAGGCTGTGGAGAGGAGC 1695
QY 1457 TGAGGTTTCACTCTCTGCTGCTGGGACATCTCTCATCAACACCGCCAGTCTCTA 1516
DB 1696 TGAGGTTTCACTCTCTGCTGCTGGGACATCTCTCATCAACACCGCCAGTCTCTA 1755
QY 1517 GAGAGCACCCCGCTACAGAGAGCCCGGCTCTGTTGGGAATCCGCCCACTCCGCTCCCT 1576
DB 1756 GAGAGCACCCCGCTACAGAGAGCCCGGCTCTGTTGGGAATCCGCCCACTCCGCTCCCT 1815
QY 1577 GTGTCCCAATGGCTGTGCTGCTCTCAATCCAGCCTACCGCTCTCTTGGCA 1636
DB 1816 GTGTCCCAATGGCTGTGCTGCTCTCAATCCAGCCTACCGCTCTCTTGGCA 1875
QY 1637 CTTACGCGCTCCCTTCAGAGCCCGGCGCCCGCCACACCGCTGGGCAACCAACA 1696
DB 1876 CTTACGCGCTCCCTTCAGAGCCCGGCGCCCGCCACACCGCTGGGCAACCAACA 1935
QY 1697 ACACCCAGGCTACAGTGGGAGCTATATGAGGCTGAGAGCCAGGCGCCCGCTTCTGC 1756
DB 1936 ACACCCAGGCTACAGTGGGAGCTATATGAGGCTGAGAGCCAGGCGCCCGCTTCTGC 1995
QY 1757 CCGCACTCCCAAGACAGCGTCCCGCTATATGCGAGGCTACATTTACCTTCAGG 1816
DB 1996 CCGCACTCCCAAGACAGCGTCCCGCTATATGCGAGGCTACATTTACCTTCAGG 2055
QY 1817 GCGTACCGGGGCAACACTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1876
DB 2056 GCGTACCGGGGCAACACTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2115
QY 1877 GCGTACCGGGGCAACACTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1936
DB 2116 GCGTACCGGGGCAACACTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2175
QY 1937 GCGAGTTTGGGAGTGCACCTGTGTGAGGTGAGAGCCCTCAAGATCTTGGTACGTTG 1996
DB 2176 GCGAGTTTGGGAGTGCACCTGTGTGAGGTGAGAGCCCTCAAGATCTTGGTACGTTG 2235
QY 1997 ATTTCCCGCTTATGTGCGTAAAGGACACCTTGTGCTGCTGCTGCTGCTGCTGCTG 2056
DB 2236 ATTTCCCGCTTATGTGCGTAAAGGACACCTTGTGCTGCTGCTGCTGCTGCTGCTG 2295
QY 2057 CAGATGCCACCAAGATG-----CCAGGAATGATTTCTTCTGAAAGAGG 2098
DB 2296 CAGATGCCACCAAGATG-----CCAGGAATGATTTCTTCTGAAAGAGG 2355
QY 2099 TGAAGATCATGTGAGGCTCAAGGACCCCAACATCATTTGGGCTGCTGGGCTGTGTGTC 2158
DB 2356 TGAAGATCATGTGAGGCTCAAGGACCCCAACATCATTTGGGCTGCTGGGCTGTGTGTC 2415
QY 2159 AGGACGACCCCTCTGCTGATGATTTACTGACTACATGAGAGACGCGGACCTCAACAGTTCC 2218

SEQUENCE CHARACTERISTICS:

LENGTH: 1197 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-445-640-7

Query Match 32.9%; Score 1197; DB 2; Length 1197;

Best Local Similarity 100.0%; Pred. No. 1.3e-256;

Matches 1197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 136 GATGCTGACATGAGGACATTTTATCCTGCCAAGTGGCCGTATGCGCTGGGCGATGCAG 195
DB 1 GATGCTGACATGAGGACATTTTATCCTGCCAAGTGGCCGTATGCGCTGGGCGATGCAG 60
QY 196 GACGGGACATCCACAGACAGTGACATCTCTCTCCAGCTCCCTGGTCCAGATTCACATGCC 255
DB 61 GACGGGACATCCACAGACAGTGACATCTCTCTCCAGCTCCCTGGTCCAGATTCACATGCC 120
QY 256 GCGCCGCCACAGAGTTTGGAGAGCAGTGACGGGATGGGCGCTGGTCCCGCGCAGGGGTGG 315
DB 121 GCGCCGCCACAGAGTTTGGAGAGCAGTGACGGGATGGGCGCTGGTCCCGCGCAGGGGTGG 180
QY 316 GTGTTTCCCAAGGAGAGGAGTACTTTCAGGTGATCTACAACTGACACTGGCTGGTGGCT 375
DB 181 GTGTTTCCCAAGGAGGAGTACTTTCAGGTGATCTACAACTGACACTGGCTGGTGGCT 240
QY 376 CTGTTGGGACCCAGGAGGACGATCCCGGGGCGCTGGGCAAGGAGTCTCCCGGAGCTAC 435
DB 241 CTGTTGGGACCCAGGAGGACGATCCCGGGGCGCTGGGCAAGGAGTCTCCCGGAGCTAC 300
QY 436 CGGCTGCGTTACTCCCGGATGGTCCGCGCTGGATGGGCTGGAAGGACCGGTGGGGTTCAG 495
DB 301 CGGCTGCGTTACTCCCGGATGGTCCGCGCTGGATGGGCTGGAAGGACCGGTGGGGTTCAG 360
QY 496 GAGGTGATCTCAGGCAATGAGACCGCTGAGGAGTGGTGTGAAGGACCTTGGGCCCGCCC 555
DB 361 GAGGTGATCTCAGGCAATGAGACCGCTGAGGAGTGGTGTGAAGGACCTTGGGCCCGCCC 420
QY 556 ATGGTTGGCCGACTGTTTCGCTTACCCCGGGCTGACCGGGTCATGAGGCTGTGCTG 615
DB 421 ATGGTTGGCCGACTGTTTCGCTTACCCCGGGCTGACCGGGTCATGAGGCTGTGCTG 480
QY 616 CGGTTAGAGCTCTATGGCTGCTTGGAGGATGAGTCTTCTTACCGCGCCCTGTG 675
DB 481 CGGTTAGAGCTCTATGGCTGCTTGGAGGATGAGTCTTCTTACCGCGCCCTGTG 540
QY 676 GGGCAGACATGTTATCTGAGGCGCTGTACCTCAAGACTCCACTATGACGGACAT 735
DB 541 GGGCAGACATGTTATCTGAGGCGCTGTACCTCAAGACTCCACTATGACGGACAT 600
QY 736 ACCGTGGGCGGACTGACATGATGGGGTCTGGCCAGCTGGCAGATGTTGTTGGGGCTG 795
DB 601 ACCGTGGGCGGACTGACATGATGGGGTCTGGCCAGCTGGCAGATGTTGTTGGGGCTG 660
QY 796 GATGACTTTAGGAAGAGTCAGGAGTGGGGTCTGGCCAGGCTATGACTATGTTGGGATGG 855
DB 661 GATGACTTTAGGAAGAGTCAGGAGTGGGGTCTGGCCAGGCTATGACTATGTTGGGATGG 720
QY 856 AGCAACACAGCTTCTCAGTGGCTATGTGAGATGGAGTTGAGTTGACCGGCTGAGG 915
DB 721 AGCAACACAGCTTCTCAGTGGCTATGTGAGATGGAGTTGAGTTGACCGGCTGAGG 780
QY 916 GCTTCCAGGCTATCAGGTCACATGTAACACATGCACAGCTGGGAGCCGCTGCTGCT 975
DB 781 GCTTCCAGGCTATCAGGTCACATGTAACACATGCACAGCTGGGAGCCGCTGCTGCT 840
QY 976 GCGGGGTGGAATGTGCTTCCGGGCTGGCCCTGCCATGGCTGGGAGGGGAGCCCATG 1035
DB 841 GCGGGGTGGAATGTGCTTCCGGGCTGGCCCTGCCATGGCTGGGAGGGGAGCCCATG 900
QY 1036 GCGCCACACCTAGGGGGCAACCTGGGGGAGCCCGAGAGCCCGGGCTGTCTCAGTGGCCCTT 1095

DB 901 CGCCACACCTAGGGGGCAACCTGGGGAGACCCAGAGCCCGGGCTGTCTCAGTGGCCCTT 960
QY 1096 GCGGGCGGTGTGGCTCGCTTTCTGCAGTGGCGCTTCTCTTTGGGGGCGCTGTGTTACTC 1155
DB 961 GCGGGCGGTGTGGCTCGCTTTCTGCAGTGGCGCTTCTCTTTGGGGGCGCTGTGTTACTC 1020
QY 1156 TTCAGCGAAATCTCTTTCATCTCTGATGTGTGAACAATCTCTCTCCGGCAGTGGGAGGC 1215
DB 1021 TTCAGCGAAATCTCTTTCATCTCTGATGTGTGAACAATCTCTCTCCGGCAGTGGGAGGC 1080
QY 1216 ACCTTCCCGCAGCGCCCTGTGGCGGCTGGCCACCTCCCAACACTTCAGCAGCTTG 1275
DB 1081 ACCTTCCCGCAGCGCCCTGTGGCGGCTGGCCACCTCCCAACACTTCAGCAGCTTG 1140
QY 1276 GAGCTGGAGCCCGAGAGGCGCAGCAGCCGCTGGCCAAAGCCCGAGGGAGCCCGACCGCC 1332
DB 1141 GAGCTGGAGCCCGAGAGGCGCAGCAGCCGCTGGCCAAAGCCCGAGGGAGCCCGACCGCC 1197

RESULT 4
US-08-456-647B-19
Sequence 19, Application US/08456647B
Patent No. 581516
GENERAL INFORMATION:
APPLICANT: Lemke Ph.D. et al., Greg E.
TITLE OF INVENTION: PROTEIN-TYROSINE KINASE GENES
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/456.647B
FILING DATE: 02-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/237,401
FILING DATE: 02-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/884,486
FILING DATE: 15-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Wetherell Ph.D., John R.
REGISTRATION NUMBER: 31,678
REFERENCE/DOCKET NUMBER: 07251/007002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 678-5070
TELEFAX: (619) 678-5099
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 3120 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
IMMEDIATE SOURCE:
CLONE: Tyro-10
FEATURE:
NAME/KEY: CDS
LOCATION: 485..3047
US-08-456-647B-19

Query Match

18.4%; Score 669.4; DB 3; Length 3120;

Best Local Similarity 56.9%; Pred. No. 1.4e-139;
Matches 1553; Conservative 0; Mismatches 986; Indels 192; Gaps 10;

Best Local Similarity 56.9%; Pred.: No. 1.4e-139;
Matches 1553; Conservative 0; Mismatches 986; Indels 192; Gaps .10;

109	CTGCTGTCGTCTTGGTGGCAAGTGAGATGCTGACATGAAGGACATTTTGTATCTCGC	168
110		
509	CTGTGTCGTCTCTGTCTTGCTATCCTCGGTTCTGCAAAAGCTCAGTTATTCAGCC	568
159	AAGTGGCCCTATGCCCTGGCATGACGAGACCGACCATCCAGACAGTACATCTCTGT	228
569	ATATGCCGCTATCTCTCTGGCGATGTCCAGGAGCCACATTTCCAGATGAGGACATCAGCC	628
229	TCCAGCTCCTGGTCAGATCCACTCGCCGCCACACAGCGTTGGAGACGATGACGGG	288
629	TCAAGTCAGTGGTCAGAACTCCAGCGCTGCCAAATATGGAGGCTGGACTCTGAAGAAGGA	688
289	GATGGGCGCTGGTCCCGCAGCGTTCGTTTCCCAAGGA--GGAGGAGTACTTTGCAG	345
589	GATGGAGCCTGGTGCTCTGAGATTCAGTGCACCCGATGACCTGAAGGAATTTCTGCAG	748
346	GTGATCTACACGACTGCACCTGGTGGCTCTGTGTGGGACCCAGGACCGCATGCCGG	405
749	ATTGACTTGGGAACCTTACACTTTATCACTCTGTGGGGACCCAGGCGCCATGCAGG	808
406	GGCTGGGCAAGGAGTTCTCCGCGACTACCGGTGCGTTACTCCCGGATGTCCGCC	465
809	GGTCATGGCATGAAATTTGACCCATGTACAAGATCAACTACAGTCGGGATGCGACTGC	868
466	TGGATGGCTGGGAAGGACCGCTGGGTCTCAGGAGTGTATCTCAGGCAATGAGGACCTGAG	525
869	TGGATCTCTGGCGTAACCGGCATGGGAAGCAGGTGCTTGATGAAACACTACCCCTAT	928
526	GAATGGTGTGAAGGACCTTGGGCCCCCATGTGTGGCCGACCTGGTTCCTTACCCC	585
929	GATGATTTCTTGAGGACTTGGAGCACCCATCTGTCGCCAGATTTGTTCGCCCTTATCCCA	988
586	CGGCTCAGCCGGTCAATGAGCGTCTCTGCGGTGAGCTCTATGCTTCGCTTCCTGGAG	645
989	GTCACGTACCACTCCATGAACGTGTGCATGAGGTTGAGCTTTATGTTGTCTGGCTA	1048
646	GATGGACTCTGTCTTACACCGCCCTGTGGGGCAGACAATGTATTATCTAGG-----	701
1049	GATGGCTTGTATCCTTACAATGCTCCAGCTGGACAGAGTTGTACTCCCTGGAGGCTCC	1108
701	-CCGTGTACTCAACGACTCCACTATACGAGACATACCGTGGCGGAGTGCAGTATGG	759
1109	ATCATTTTATCTGAATGATTTCTGTATGATGAGCTGTTGGGTACAGCATGACTGAAGG	1168
760	GGTCTGGCCAGCTGGCAGATGGTGTGTGGGCTGATGACTTTAGGAAGAGTCAAGAG	819
1169	---CTAGGCCAGTTGACTGATGAGATATCCGCGCTGGATGATTTTACCCAGACCCATGA	1225
820	CTGCGGTCTGGCCAGGCTATGACTATGTGGATGGAGCAACACAGCTTCTCCAGTGC	879
1226	TACCAGCTGTGGCCTGGCTATGACTACGTGGATGGCGGAACGAAAGTGCATCCACAGT	1285
880	TATGTGAGATGGAGTTTGAGTTTACCGCGCTGAGGCGCTTCCAGGCTATGCAGGTCCAC	939
1286	TTCATTCAGATCATGTTTGAATTTGACCGCAATCAGGAATTTTACTACATGAAGGTCCAC	1345
940	TGTAACACATGCACAGCTGGGAGCCGCTGTCGCTGGCGGGTGGAAATGTCGCTTCCGG	999
1346	TGCAACAACATGTTTCTAAAGGTGTGAAGATTTTAAAGAGGTCCAGTGTCTACTTTCGC	1405
1000	CGTGGCCCTGCCATGCTCTGGGAGGGGAGCCCATGCGCCACACACCTTAGGGGCAACCTG	1059
1406	TCGG---AAGCCAGAGTGGGAACCCACTGCTGTCTACTTTTCCCTTGGCTCTGGACGAT	1462
1060	GGGAGCCCCAGAGCCCGGCTGTCTCAGTGGCCCTTGGCGGCCGCTGTGGCTTCCTTCG	1119
1463	GTGAACCCAGTGCCTGGTTGTACGGTGGCCCTCCACACACGAAATGGCCAGTGCCTC	1522
1120	CAGTGGCGCTTCTCTTTTGGGGGCCCTGGTTACTCTTCTACGCAAAATCTCCTTCACTCT	1179

Db	1523	AAGTGC	AATACCATTTTCCGACACGTGGATGATGTTTACGCGAGATCACCATTCTTCCAATCA	1582
Qy	1180	GATGTG	TGAACAATTCCTCTCCGGGCACTGGGAGCAGCACTTCCCGCCAGCCCTTGGTGG	1239
Db	1583	GATGCT	GCATGTATAAACAACCTCTGGAGCCCTTCCCACTCTCCTA-----	1629
Qy	1240	CCGCTG	GGCCCCAOCCTCCCACCAACTTCAGACGTTGGAGCTGGAGCCAGAGCCACGAG	1299
Db	1629	-----	-----TGGCACCCACCACCATATGAT	1648
Qy	1300	CCGCTG	GCCAAGCCCGAGGGAGCCCGACCGCCATCTCTCATCGCTGCCTGGTGCCCATC	1359
Db	1649	CCCATG	CTTAAAGTTGATAGCAACACATCGGATTCCTGATTTGGTTGGTGCCCATC	1708
Qy	1360	ATCTGCT	CTCTGCTCAATATGGCCCTCATGCTCTGGCGGCTGACGTGCGCAGGCTC	1419
Db	1709	ATCTTAT	CTCTGCTATCATCTGTCATCTCTGGAGGCAGTTCTGGCAGAGATG	1768
Qy	1420	CTCAGAA	AGGCTGAACGGAGGGTGTGGAGAGGAGCTGACGGTTTACCTCTCTGTCCCT	1479
Db	1769	CTAGAAA	AGGCTTACGGAGGATGCTGGATGATGAATGACAGTCAAGCTTTCCTCGCC	1828
Qy	1480	GGGACAC	TATCTCATCAACAACGCCCGCAGTCTAGAGAGCCACCCCTTACAGGAG	1539
Db	1829	AGCGAG	TCCAGCATGTTTAAATAACAACCGCTCCTC-----ATCACCAAGTGAACAGGAG	1882
Qy	1540	CCCGGG	CGCTCGTGGGAATCCGCCCACTCCGCTCCCTGTGTGCCCAATGGCTCTGCGTTG	1599
Db	1883	TCCAACT	CTACTTATGATCGAATCTTCCCTTCCCTGCTGACTACGAGAGCCATCCAGA	1942
Qy	1600	CTGCTCT	TCCAATCCAGCCTACCGCCCTCTTCTGGCCACTTACGCCCGTCCCTCGAGGC	1659
Db	1943	CTGATC	-----	1949
Qy	1660	CCGGG	CCCCCCCCACCCGCTTGGGCCAAACCCACCAACCCAGGCCCTACAGTGGGGAC	1719
Db	1949	-----	-----CGAAGTCTCCAGATTTGCTCCAGGAGAGGAGA	1983
Qy	1720	TATATG	GAGCCTGAGAAGCCAGGCGCCCGCTTCTGCCCCACCTCCCCAGAACAGGCTC	1779
Db	1984	GTCA	GGTGCAGTGTGTGTGAGCCGGCC-----CAGCCCAATGGACCTTAGGGCGTG	2038
Qy	1780	CCCCAT	TATCCGAGGCTGACATTTTACCTTCGAGGCGCTCACCGGGGGCAACACCTAT	1839
Db	2039	CCCCACT	TATCAGAAGCCGACATAGTGAATCTCCAGGAGTGACAGTGGCAACACCTAC	2098
Qy	1840	GCTGTG	CGCTCACTGCCCCCG-----GGGCGTGGGGATGGGCCCCCAGAGTGGATTTC	1896
Db	2099	TGTGTG	CCCTGCTGCTAACCATTGATCTGCTATCGGGGAAGATGTGGCTGTGGAAGATTCT	2158
Qy	1897	CCTCGAT	CTCGACTCCGCTTCAAGGAGAGCTTGGCAGGCGCCAGTTTGGGAGGTGCAC	1956
Db	2159	CCAGGAA	CTGTGGCCCTTCAAGGAGAAGCTTGGGAGAGCCAGTTTGGGAGGTTCACT	2218
Qy	1957	CTGTGT	GAGTTCAGACGCCCTCAAGATCTGTGTCAGTCTTGATTTCCCTCTTAATGTGCGT	2016
Db	2219	CTCTGT	GAGTGGAGGAATGGAAAAATTCAAAGACAAAGATTTTGCATAGATGTCAGT	2278
Qy	2017	AAGGGAC	ACCCCTTTGCTGGTACTGTCAAGATCTTACGGCCAGATGCCACCAAGAATGCC	2076
Db	2279	GCCAAAC	CGCCTCTCTGGTGGCCGTGAAAAATGCTCCGAGCAGATGCCAACAGAATGCC	2338
Qy	2077	AGGAAT	GTATTTCTGAAAGAGGTGAAGATCATGTGCGAGGCTCAAGGACCCAAACATCAT	2136
Db	2339	AGGAAT	GTATTTCTTAAGAGATCAAGATCATGTCTCGGCTCAAGGACCCAAACATCATC	2398
Qy	2137	CGGCTG	CTGGCGGTGTGTGCGAGGACGACCCCTCTCGATGATTACTGACTACATCGAG	2196
Db	2399	CGTCTCT	TAGTGTGTGCATCACTGAGGACCCGCTCTCGATGATCAAGATATCATCGAG	2458
Qy	2197	AACGGCG	ACCTTCAACAGTTCTTCAAGTGGCCACCAAGCTTGGAGGACAAGGACCGAGGGG	2256
Db	2459	AATGGAG	ATCTTAAATCAGTTTCTTTCTCGCCACGAGCTCTGA-----	2502

QY 2257 GCGCTGGGACGGGAGGCTGGCGAGGGGCCACCATCAGCTACCCAAATGCTGTCAT 2316
Db 2502 -----GTTCTCTTCTAGTGATGCCACATCAGTACGTCAGTAACTGAGTTT 2548
QY 2317 GTGGCAGCCAGATCGCTCCGCGATCGGTATCTGGCCACACCTCACTTTGTACATCG 2376
Db 2549 ATGGCAACCCAGATGCTGCTGATGATGAGTACCTTTCTCTCACTTTGTCCACCGA 2608
QY 2377 GACCTGCCACGGGAACTCCCTAGTGTGGGAAATTTCAACATCAAAATCGCAGACTTT 2436
Db 2609 GATCTGCCACAGAACTCTTTAGTGGCAAGAAATACACCATCAAGATAGTGTATTT 2668
QY 2437 GGCATGAGCCGGAACCTCTATGCTGGGACTATTTACCGTGTGCGAGGGCCGGCAGTGTG 2496
Db 2669 GGCATGAGCAAAACCTGTACAGTGTGATTTACTACCGGATCCAGGGCCGGGGTGTCTC 2728
QY 2497 CCATTCGCTGGATGGCTGGGAGTGCATCTCATGGGAAATTCACGACTCGGAGTGAC 2556
Db 2729 CCCATTCGCTGGATGCTCTGGGAAAGCATCTTCTGGCAAAATTCACCAAGCAAGTGTAT 2788
QY 2557 GTGTGGCCCTTTGGTGTGACCCCTGTGGGAGGTGCTGATGCTCTGTAGGGCCCGCCCTTT 2616
Db 2789 GTGTGGCCCTTTGGTGTGACCTGTGGGAGACCTTCACCTTTTGGCAGGACAGCCCTAT 2848
QY 2617 GGGCAGCTCACCGACGAGCAGGTATCTGAGAACGCGGGGAGTCTTCCGGGACCAAGGGC 2676
Db 2849 TCCAGCTGTGGATGAGCAGGTATATCGAGAACACTGGAGAGTCTTCCGAGACCAAGGG 2908
QY 2677 CGGAGCTGTACCTGTCCCGCCGCTGCTGCGGAGGAGTCTTCCCGGAGGAGTGTATG 2736
Db 2909 AGGAGATCTATCTCCCTCAACAGCCCTTTGGCCGAGCTCTGTGTATAGCTGTATGCTC 2968
QY 2737 CGGTGTGGAGCGGGAGTCTGAGCAGCAGCAGCCCTTTTCCAGCTGCTGCTGCTGCTG 2796
Db 2969 AGCTGTGGAGAGAGAACCAAGCAGCCGCGCATCTTCCAGGAATACACCTCTCTGCTT 3028
QY 2797 GCAGAGATGACTCAACACGGTGTGAATCA 2827
Db 3029 CTTGAGCAAGGAGCGAGTGATGATGATCA 3059

RESULT 5

US-08-237-401A-19
; Sequence 19, Application US/08237401A
; Patent No. 5837448
; GENERAL INFORMATION:
; APPLICANT: Lemke Ph.D. et al., Greg E.
; TITLE OF INVENTION: PROTEIN-TYROSINE KINASE GENES
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/237,401A
; FILING DATE: 02-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/884,486
; FILING DATE: 15-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile Ph.D., Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07251/007001

TELECOMMUNICATION INFORMATION:

; TELEPHONE: (619) 678-5070
; TELEFAX: (619) 678-5099
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3120 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; IMMEDIATE SOURCE:
; CLONE: Tyro-10
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 485..3047
US-08-237-401A-19

Query Match 18.4%; Score 669.4; DB 3; Length 3120;
Best Local Similarity 56.9%; Pred. No. 1.4e-139;
Matches 1553; Conservative 0; Mismatches 986; Indels 192; Gaps 10;

QY 109 CTGCTGCTGCTCTTGGTGGCAAGTGGAGATGCTGACATGAAGGACATTTTGTATCTCTGCC 168
Db 509 CTGGTGTCTGCT 568
QY 169 AAGTGGCCCTATGCGCTGGGATGCGAGGACCGGACCATCCAGACAGTACATCTCTCTGCT 228
Db 569 ATATGGCCCTATCTCTCTGGGATGCTGAGGAGGCCACATTCAGATGAGGACATCACAGCC 628
QY 229 TCCAGCTCTCTGTCAGATTCACATGCGCGCCGCCACAGCAGGTTGGAGAGCAGTACGCGG 288
Db 629 TCAAGTCACTGCTGAGATCCAGGCTGCCAAATATGGAGGCTGGACTCTGAAGAAGGA 688
QY 289 GATGGGCTGTGCTGCGCCGAGGTGCTGCTTCCCAAGGA---GGAGGAGTACTTTCAG 345
Db 689 GATGGAGCTGTGCTCTGAGATTCAGTGCACCCATGACCTGAGGATTTCTGCGAG 748
QY 346 GTGATCTTACACGACTGCACCTGCTGGCTCTGCTGGGACCCAGGACCGCATGCGCGG 405
Db 749 ATTGACTTTCGCAACCTACACTTTTATCACTCTTGTGGGAGCCAGGCGGCGCATGCGAGG 808
QY 406 GGCCTGGCAAGGAGTCTCCCGGAGCTACCGGCTGCTTACTCCCGGATGCTGCGCGC 465
Db 809 GGTCACTGCAATGAAATTTGCACCATGTACAAGATCAACTACAGTCGGGATGGCAGTCG 868
QY 466 TGGATGGCTGTGAAGGACCGCTGGGTCAGGAGGTGATCTCAGGCAATGAGGACCTGAG 525
Db 869 TGGATCTCTGCGTAAACCGCATGGGAGCAGGTGCTTGTGGAACAGTAAACCTTAT 928
QY 526 GGAGTGTGTGAAGGACCTTGGGCCCCCATGTTGCGGACCTGCTGCTGCTTCTACCCC 585
Db 929 GATGTATTCTCTGAAGGACTTGGAGCCACCATCTGCGCAGATTTGTGCGCTTATCCCCA 988
QY 586 CGGCTGACCGGGTCATGAGGCTGTCTGCGGTCAGGCTCTATGCTGCTGCTGAGG 645
Db 989 GTCAGTGACCACTCCATGAACGTGTGATGAGGTTGAGCTTTATGTTGTGTCTGCTGA 1048
QY 646 GATGACTCTCTGCTTACACCCCGCTCTGGGCGACAGCAATGATTTATCTGAGG----- 701
Db 1049 GATGGCTTGGTATCCTACAATGCTCCAGCTGGACAGCAGTTTGTACTCCCTGGAGGCTCC 1108
QY 701 -CCGTGTACCTCAAGACTCCACCTATGACGACATACCCGTGGGCGGACTGCAATATGGG 759
Db 1109 ATCAATTTATCTGAATGATTTCTCTATGATGAGCTGTGGGTACAGCATGACTGACGAGG 1168
QY 760 GGTCTGGCCAGCTGGCAGATGTTGTTGGGCTGGATGACTTTAGGAAGAGTCAGAG 819
Db 1169 ---CTAGGCGAGTTGACTGATGGAGTATCCGCGCTGGATGATTTACCCAGACCCATGAA 1225
QY 820 CTGCGGGTCTGGCCAGGCTATGACTATGTGGGATGGACCAACACAGCTTCTCCAGTGGC 879
Db 1226 TACCACGTGTGCGCTGCTATGACTACGTGCTGGTGGGCGGAGGAAAGTGTCTACCAACGT 1285

QY 880 TATGTGAGATGAGATTTGAGTTTGACCGCTGAGGGCCCTCCAGGCTATGAGGTCCAC 939
DB 1286 TTCAATGAGATCATGTTTGAATTTGACCGAATCAGGAATTTTACCATGAGGTCCAC 1345
QY 940 TGTAAACAATGACACACGCTGGGAGCCCTCTGCTGCGGGGGTGAATGTGCTTCCGG 999
DB 1346 TGCACAACATGTTGCTAAAGGTGTGAAGATTTTAAAGAGGTCCAGTCTACTTTCGC 1405
QY 1000 CGTGGCCCTGCGCATGCGCTGGAGGGGAGCCCATGCGCCACAACCTAGGGGCAACCTG 1059
DB 1406 TCGG---AAGCCAGCAGTGGGAACCACTGCTGTACTTCTTCCCTGCTCTGACCAT 1462
QY 1060 GGGGACCCAGAGCCGCGCTGTCTCAGTGGCCCTTGGGGCCGTGGCTGCTGCTTCTG 1119
DB 1463 GTGAACCCAGTGGCGGTTGTACGGTGGCCCTCCACCAAGTGGCCAGTGGCCATC 1522
QY 1120 CAGTGGCGCTTCTCTTTTGGGGGGCCCTGTTACTTCTCAGGGAATCTCTTTCATCTCT 1179
DB 1523 AAGTGCCAATACCATTTTCCGACAGGTGGATGATGTTTCCAGGAGATCACITTTCCAATCA 1582
QY 1180 GATGGGTGAACAATTCCTCTCCGGACTGGAGGACCTTCCCGCCAGCCCTGCTGTGG 1239
DB 1583 GATGCTGCAATGTATAACAATCTGGAGCCCTTCCACCTCTCTTA----- 1629
QY 1240 CCGCTGGCCCACTCCCACTTACAGCTTGGAGCTGGAGCCCAAGAGCCAGGAG 1299
DB 1629 -----TGGCACCCACCACTATGAT 1648
QY 1300 CCGTGGCAAGCCGAGGGAGCCGACCGCATCTCATCGGCTGCTGTGGCCATC 1359
DB 1649 CCATGCTTAAAGTTGATGATAGCAACTCGATCTCTGATGGTGTGGTGGCCATC 1708
QY 1360 ATCTGCTCTCTGCTCATATGCTTCCCTCATGCTTGGCGCTGCACTGGCGAGGCTC 1419
DB 1709 ATCTTCTCTGCTGCTATCATCTCATCTGTGGAGCACTTCTTGGCAGAAGATG 1768
QY 1420 CTCAGCAAGCTGACGAGGCTGTGGAGAGGAGCTTCCAGGTTTCACTCTCTGCTCCCT 1479
DB 1769 CTAGAAAGCTTACGAGGAGTGTGATGATGAATGACAGTCAAGCTTTCCTGCTCC 1828
QY 1480 GGGGACATATCTCATCAACAACCGCCAGGTCTTAGAGAGCCACCCCTTACCAGGAG 1539
DB 1829 AGCAGTCCAGCATGTTCAATAACAACCGCTCTC-----ATCACAAGTGAACAGGAG 1882
QY 1540 CCGCGGCTCTGGGAATCCGCCCCACCTCCGCTCCCTGTGTCCCAATGGCTTGCCTTG 1599
DB 1883 TCCAACCTTACTTATGATCGAATCTTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1942
QY 1600 CTGCTCTCAATCCAGCTTACCGCTCTCTTGGCCACTTACGCCCTGCTGAGGC 1659
DB 1943 CTGATC----- 1949
QY 1660 CCGGGCCCCCACACCGCCCTGGGCCAAACCCACACCCAGGCTTACAGTGGGGAC 1719
DB 1949 -----CGAAGCTTCCAGATTTGCTCCAGGAGGAGGA 1983
QY 1720 TATATGGAGCTGGAAGCCAGGCGCCCGTCTTGTCCCACTTCCCGGAGAGGCTC 1779
DB 1984 GTCAGGCTGAGTGTGTGTGAAGCCGCT-----CAGCCCAATGGACCTGAGGGGCTG 2038
QY 1780 CCCCATTATCGGAGCTGACATTTTACCTGAGGCGGCTCACGGGGGCAACACCTAT 1839
DB 2039 CCCCATTATGAGAGCCGACATGATGATGCTGCTATCGGGGAAGATGTGGCTGTGGAAGCTT 2158
QY 1840 GCTGTGCTGCACTCCCGCCAG---GGCAGTGGGGATGGCCCCCCCCCAGAGTGTGATTC 1896
DB 2099 TGTGTGCTGCTGTAAACCATGATGCTGCTATCGGGGAAGATGTGGCTGTGGAAGCTT 2158
QY 1897 CTTCAATCTGCACTCCGCTTCAAGAGAGCTTGGCGAGGCGGCTTGGGGAGGCTGAC 1956
DB 2159 CCGAGGAACCTGTTGGCTTCAAGAGAGAGCTGGGAGAGCCAGTGTGGGGAGGTTTCA 2218

QY 1957 CTGTGTGAGTTCGACAGCCCTCAAGATCTGGTCACTTGTGATTTCCCTTAAATGTGCGT 2016
DB 2219 CTCTGTGAAGTGGGAATGGAATAATCAAAGACAAAGATTTGCACCTAGATGTCACT 2278
QY 2017 AAGGGACACCTTTGCTGTGTAGTGTCAAGATCTTACGGCCAGATGCCACCAAGATGCC 2076
DB 2279 GCCAACCAGCTCTCTCTGGTGGCGGTGAAATGCTCCGAGCAGATGCCAACAAGATGCC 2338
QY 2077 AGGAATGATTTCTTGAAGAGGTGAAGATCATGTGAGGCTCAAGGACCAACACATCAT 2136
DB 2339 AGGAATGATTTCTTGAAGAGATCAAGATCATGTCTGGCTCAAGGACCAACACATCAT 2398
QY 2137 CGGTGCTGGGCTGTGTGTGAGGAGACCCCTCTGCTGATGATTTACTGACTACATGGAG 2196
DB 2399 CGTCTCTTAGTGTGTGATCATCACTCACTGAGGACCGCTCTGCTGATGATCAAGATGAG 2458
QY 2197 AAGGGACACCTCAACAGTTCTCTGAGTGGCCACCACTGAGGAGCAAGGACCCGAGGG 2256
DB 2459 AATGGAGATCTTAAATCAGTTTCTTCTGCGCACGAGCTCTGA----- 2502
QY 2257 GCGCTGGGACGCGGAGGCTGCGAGGCGCCACCATCAGCTACCAATGCTGCTGCAT 2316
DB 2502 -----GTTCTGTTCTAGTGTGATGCCACAGTCACTGACCTGAAGTTT 2548
QY 2317 GTGGACCCAGATCGCTCCGGCATCGGCTATCTGGCCACACTCAACTTTGTACATCGG 2376
DB 2549 ATGGCAACCCAGATGCTCTGCTGTGATGAAGTACCTTTCTGCTCTCACTTTGCTCCCGA 2508
QY 2377 GACTGGCCACGCGGAACTGCTAGTGGGAAATTTTACCAATCAAAATCGCAGACTTT 2436
DB 2609 GATCTGGCCACAGAACTGTTTGTGGGCAAGAAATTTACCAATCAAGATAGCTGATTT 2668
QY 2437 GGCATGAGCGGAACTCTATGCTGGGAGTATTACCGTGTGCGAGGCGCGGAGTGTG 2496
DB 2669 GGCATGAGCAAGAACTGTACAGTGTGATTAACCGGATCCAGGCGCGGCGGCTGCTC 2728
QY 2497 CCATCGCTGAGTGGCTGGGAGTGCATCTCATGGGAAGTTTACGACTTGCAGTGTGAC 2556
DB 2729 CCATTCGCTGAGTGTCTGGGAAAGCATCTTGTGGGCAAAATTTACCAAGTCAAGTAT 2788
QY 2557 GTGTGGCTTTGCTGTGAGCCCTGTGGGAGTGTGATGCTGTGAGGCGCGAGCCCTTT 2616
DB 2789 GTGTGGCTTTTGGGTGACTCTGTGGGAGACTTTCACCTTTTGGCAGGAGCAGCCCTAT 2848
QY 2617 GGCAGCTCACCCAGCAGCAGGCTCATCGAGACCGGGGAGTCTTCCGGACCAAGGCG 2676
DB 2849 TCCAGCTGTGGATGAGCAGTTCAGAGACACTTGGAGAGTCTTCCGAGACCAAGGG 2908
QY 2677 CGGAGGTGTACCTGTCCCGGCGCTGCTGCTCCGAGGCGCTATATGAGTGTGCTT 2736
DB 2909 AGCAGATCTATCTCCCTCAACCAAGCCCTTTGCGCGACTCTGTGTATAGTGTGCTC 2968
QY 2737 CGTGTGTGGGCGGAGTCTGAGGAGCAACCCCTTTTCCAGCTGATCGTTCCTG 2796
DB 2969 AGCTGTGGAGAGAGAAACCAAGCACCAGCCGCTCTTCCAGGAAATACACCTCTGCTT 3028
QY 2797 GCAGAGATGCTCACTCAACCGGTGTGAATCA 2827
DB 3029 CTTACAGGAGGCGCGGATGATGATCA 3059

RESULT 6

US-08-336-343A-3

; Sequence 3, Application US/08336343A

; Patent No. 5677144

; GENERAL INFORMATION:

; APPLICANT: Ullrich, Axel

; APPLICANT: Alves, Frauke

; TITLE OF INVENTION: CCK-2, A No. 5677144el Receptor Tyrosine Kinase

; NUMBER OF SEQUENCES: 43

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pennie & Edmonds

; STREET: 1155 Avenue of the Americas

CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/336,343A
FILING DATE: 08-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7683-065
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3157 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 370..2934
US-08-336-343A-3

Query Match 18.38; Score 666.8; DB 2; Length 3157;
Best Local Similarity 56.88; Pred. No. 5.3e-139;
Matches 1532; Conservative 0; Mismatches 977; Indels 189; Gaps 9;

QY 110 TGCTCTGCTCTTGGTGGCAAGTGAGATGCTGACATGAAGGGACATTTTATCCTGCCA 169
DB 395 TGGTCTGCTCTGCTGCTGCTATCTTGAGTTCTGCAAAAGCTCAGTTAATCCACTA 454
QY 170 AGTGGCGGTATGCCCTGGGATGACAGGACCGACCATCCAGACAGTGCATCTCTCTT 229
DB 455 TATGCCGTATCTCTGGGATGTCAGGAGGCCAGATTCAGATGAGGACATCACAGCTT 514
QY 230 CCAGCTCTGCTCAGATTCACCTGCGCCCGCCACAGAGCTTGGAGAGCAGTGACCGGG 289
DB 515 CCAGTCACTGGTCAAGTCCAGCTGCGCAATATGGAAGGCTGGACTCAAGAAAGGGG 574
QY 290 ATGGGGCTGTGTCGCCCGCAGGGTGGTGTTCCTCAAGGA---GGAGGAGTACTTGCAGG 346
DB 575 ATGGAGCTGTGGCTCCTGAGATTCAGTGGACCTGATGACCTGAAGGAGTTTCTGCAGA 634
QY 347 TGGATCTACAAGCACTGACCTGCTGCTGCTGGTGGGACCCAGGACGCGCATGCCGGGG 406
DB 635 TTGACTTGCACACCTCCATTTATCACTCTGTTGGGAGCCAGGGCGCCATGCAGGAG 694
QY 407 GCTTGGCAAGAGTTCTCCCGAGCTACCGGCTGCTTACTCCCGGATGGTCCGCGCT 466
DB 695 GTCAATGGATCAGATTGCGCCCATGTACAAGATCAATTAAGTGGGATGGCACTCGCT 754
QY 467 GGATGGGTGGAAGGACCGCTGGGTCAGGAGGTGATCTCAGGCAATGAGGACCTCGAGG 526
DB 755 GATCTCTGGCGAACCCTCATGGAAACAGGCTGCTGGATGGAATAGTAAACCCCTATG 814
QY 527 GAGTGGTCTGAAGCACTTGGGCCCCCATGTTGGCCCGACTGCTGCTTCTACCCCC 586
DB 815 ACATTTTCTAAAGGACTTGGAGCGCCCATTTAGCCAGATTTTTCGGTTCATTCAG 874
QY 587 GGGCTGACCGGGTCAATGAGGCTCTGCTGCGGGTAGAGCTCTATGGCTGCTCTGGAGG 646

DB 875 TCACCGACCACTCCATGAATGTGTGATGAGAGTGGAGCTTACGGCTGTCTGGCTAG 934
QY 647 ATGGACTCTGTCTTACACCGCCCTGTGGGCGACAGCAATCTATTTATCTGAGG----- 701
DB 935 ATGGCTTGGTGTCTTACAATGCTCCAGCTGGCAGCAGTTTGTACTCCCTGGAGTTCCA 994
QY 701 CGGTGTACTCTAAGACTTCCACCTATGACGACATACCGTGGGCGGACTGCAAGTATGGGG 760
DB 995 TCATTTATCTGAATGATTCTGTCTATGATGAGCTGTGTGGATACAGCATGACAGAAAGG- 1054
QY 761 GTCTGGCCAGCTGGCAGATGTTGGTGGGCTGGAGTCTTTAGGAAGAGTCCAGAGC 820
DB 1054 --CTAGGCCAATTGACCATGGTGTGCTGGCTGGAGCAATTTCCACGAGACCATGAT 1111
QY 821 TCGGGTCTGGCCAGGCTATGACTATGTGGATGGAGCAACACAGCTTTCTCCAGTGGCT 880
DB 1112 ACCACGTGTGGCCCGCTATGACTATGTGGCTGGCGAAGAGAGTCCACCAATGGCT 1171
QY 881 ATGTGAGATGGAGTTTGAAGTTTACCGGCTGAGGGCTTCCAGGCTATGCAAGTCCACT 940
DB 1172 ACATTTGAGATCATGTTTGAATTTGACCGCATCAGCAATTTCACTACCATGAAGTCCACT 1231
QY 941 GTACAACATGCACACGCTGGGAGCCGCTCTGCTGGGCGGTGGAATGTCGCTTCCGCG 1000
DB 1232 GCAACAACATGTTTCTAAAGTGTGAAGATCTTTAAGAGGTACAGTCTACTTCCGCT 1291
QY 1001 GTGGCCCTGCCATGGCTTGGAGGGGAGCCCATGCGCCACAACTAGGGGCAACCTGG 1060
DB 1292 CTG---AAGCCAGTGAAGTGGGAACCTAATGCAATTTCTTCCCTTCTCTGGATGAG 1348
QY 1061 GGGACCCAGAGCCCGGCTGCTCAGTGGCCCTTGGGCGGCTGTGGCTGCTTCTGCT 1120
DB 1349 TCAACCCAGTGTCTGCTGTTGTACGGTGGCTCTCCACACCAAGATGGCCAGTGCCATCA 1408
QY 1121 AGTGGCGCTTCTCTTGGGGGCGCTGTTACTCTTTCAGGAAATCTCTTCTATCTCTG 1180
DB 1409 AGTGTCAATACCATTTGAGATGATCTGGATGATGTTTCAAGTATGATCCTTCCATCAG 1468
QY 1181 ATGTGGTGAACAATTTCTTCTGCGCACTGGGAGGACACTTCCCGCCAGCCGCTGGTGG 1240
DB 1469 ATGTGCAATGTACAACAACCTCTGAAGCCCTGCCACCTCTCTA----- 1514
QY 1241 CGCTGGCCCACTCCACCAACTTCAGCAGCTTGGAGCTGGAGCCGACAGGCGCAGC 1300
DB 1514 -----TGGACCCCAACACTATGATC 1534
QY 1301 CGGTGGCCAAAGCCCGAGGGAGCCGACCGCATCTCTATCGGCTGCTGTGGGCAATCA 1360
DB 1535 CAATGCTTAAAGTTGATGACAGCAACACTCGGATCTGATTTGGTGGTGGCAATCA 1594
QY 1361 TCTGCTCTGCTGCTCATCATTTGCCCTCATGCTCTGGGGGCTGCACTGGGCGAGGCTCC 1420
DB 1595 TCTTTATCTCTGCTGGCATCATTTGATCATCTCTGGAGGAGTCTTGGCAGAAATGC 1654
QY 1421 TCAGCAAGCTGACGGAGGGTGTGGAAGAGAGCTGACGTTTCACTCTCTGCTGCTG 1480
DB 1655 TGGAGAGGCTTCTCGGAGGATGCTGGATGATGAATGACAGTCAAGCTTTTCCCTGGCAA 1714
QY 1481 GGGACACTATCTCATCAACAACCGCCAGGTCTCTAGAGAGCCACCCCGCTTACCAAGAGC 1540
DB 1715 GTGATTTAGCATGTTCAACAT-----AACCCTCTCT 1747
QY 1541 CCGGGCTCGTGGGAATCCGCCCCACTCCGCTTCCCTGTGTCTCCCAATGGCTGTGGTGGC 1600
DB 1748 CATCACTAGTGAACAAGGGTCCAACTCGACTTACGATCGCATCTTT----- 1795
QY 1601 TGCTCTCCAATCCAGCTACCGCTTACCCTCTTCTTGCCCATACCGCCGCTCCCTCGAGGC 1660
DB 1795 -----CCCCCTCGGC 1804
QY 1661 CGGGCCCCCACCACCGCTGGGCAAAACCCACCAACCCAGGCGCTACAGTGGGGACT 1720

Db 1805 CTGACTACAGGAGCCATCCAGGCTCATAGGAAACTCCAGAAATTTGCTCCAGGGGAGG 1864
QY 1721 ATATGAGCCTGAGAACCCAGGCGCCCGCTTCTGTCGCCACCTCCCCAGAACAGCGTCC 1780
Db 1865 AGGAGTCAGGCTGAGCGGTGTTGTAAGCCAGTCAGCCAGTGCCTGAGGGGTGC 1924
QY 1781 CCATATGCGGAGGCTGACATGTTTACCTCGAGGGGTGTCAGGGGGGCAACACTATG 1840
Db 1925 CCACATGTCAGAGGCTGACATGTAAGCTCCAGGAGTGCAGGAGGCAACATATCT 1984
QY 1841 CTGTGCTGCTGACCTGCCCCAGGGGCACTC---GGGATGGGCCCCCAGAGTGGATTTC 1897
Db 1985 CAGTGCCTGCCCTCACCATTGGACCTGCTCTCAGGAAAGATGTGGTGTGGAGGATTTC 2044
QY 1898 CTCGATCTGACCTCCCTTCAAGGAGAACTTGGGAGGGCCAGTTTGGGAGGTGCACC 1957
Db 2045 CCAGGAAACTCTTCACTTTTCAAGAAAGCTGGGAGAGGAGTGGGGAGGTTTATC 2104
QY 1958 TGTGTGAGTTCGACAGCCCTCAGATCTGCTGAGTCTGATTTCCCTTATGTGGTA 2017
Db 2105 TCTGTGAAGTGGAGGAATGGAAATTCAGAACAAAGATTTGCCCTAGATGTCACTG 2164
QY 2018 AGGAGACACCTTGTGCTGAGTGTCAAGATCTTACGGCCAGATGCCACCAAGATGCCA 2077
Db 2165 CCAACAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2224
QY 2078 GGAATGATTTCTGGAAGAGGTGAAGATCATGTCGAGGCTCAAGGACCAACATCATTC 2137
Db 2225 GGAATGATTTCTTAAGGAGATAAGATCATGCTCGGCTCAAGGACCAACATCATTC 2284
QY 2138 GGTGCTGGGCTGTGTGTCAGGAGGACCCCTGCTGATGATTTACTGACTACATGGAGA 2197
Db 2285 ATCTATTATCTGTGTATCACTGATGACCCCTCTGATGATCATCTGAATACATGGAGA 2344
QY 2198 AGGCGACCTCAACCACTTCTCAGTGCACCCAGCTGAGGAGACAGGACCGAGGGG 2257
Db 2345 ATGGAGATCTCAATCAGTTCTTCCCGCAGAGCCCTTAATTTCTTCCACGCG--- 2402
QY 2258 CCCTGGGAGCGGAGGCTGCGAGGGGCCACCATCAGCTACCAATGCTGCTGATG 2317
Db 2402 -----ATGTACGCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2437
QY 2318 TGCGACCCAGATCGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2377
Db 2438 TGCTACCCAAATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2497
QY 2378 ACCTGCGCACGGGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2437
Db 2498 ATCTGCGCACAGAACTGTTAGTGGGTAGAACTACAACTCAAGATAGCTGACTTGG 2557
QY 2438 GCATGAGCGGAACCTCTATGCTGGGACTATTTACCGTGTGCGAGGGCGGCACTGCTGC 2497
Db 2558 GAATGAGCAGGAACCTGTACAGTGGTGTGCTATTTACCGGATCCAGGGCGGCACTGCTGC 2617
QY 2498 CATCGCTGAGTGGCTGGGAGTGCATCCCTCATGGGAAGTTACAGTCTGGAGTACG 2557
Db 2618 CTATCGCTGAGTGTGTTGGGAGATATCTGCTGGGCAAGTTTCACTACAGCAAGTATG 2677
QY 2558 TTGGGCGCTTTGCTGAGCCCTGCTGGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2617
Db 2678 TGTGGGCGCTTTGCTGAGTGTGCTGGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2737
QY 2618 GCGAGCTACCGCAGCAGGAGTTCATCGAGAACCGGGGAGTTCCTCCGGGACCAAGGCGCC 2677
Db 2738 CCAGCTGTGATGATGAGGTTATTGAGATATCTGGAGAGTTCCTCCGAGACCAAGGGA 2797
QY 2678 GCGAGTGTACCTGTCCTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2737
Db 2798 GCGAGACTTACCTTCCCTCAACCGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2857
QY 2738 GGTGTGTGAGCGGGAGTCTGAGCAGGACCAACCCCTTTTCCAGCTGATCGTCTCT 2795
Db 2858 GCTGCTGGAGAGATACGAGAGACCGCTCCCTCAATCCAGAAATCCACCTTCTGCT 2915

RESULT 7

US-08-336-343A-5/C
; Sequence 5, Application US/08336343A
; Patent No. 5677144
; GENERAL INFORMATION:
; APPLICANT: Ullrich, Axel
; APPLICANT: Alves, Frauke
; TITLE OF INVENTION: CK-2, A No. 5677144el Receptor Tyrosine Kinase
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/336,343A
; FILING DATE: 08-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7683-065
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3157 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-336-343A-5

Query Match 18.3%; Score 666.8; DB 2; Length 3157;
Best Local Similarity 56.8%; Pred. No. 5.3e-139;
Matches 1532; Conservative 0; Mismatches 977; Indels 189; Gaps 9;

QY 110 TGCTGCTGCTCTTGGTGGCAAGTGGAGATGCTGACATGAAGGAGCAATTTGATCTCGCA 169
Db 2763 TGGTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2704
QY 170 AGTCCGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 229
Db 2703 TATGCGCTATCTCTGCGCATGCTGAGGAGCCAGATTCAGATGAGGACATCAGACTT 2644
QY 230 CCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 289
Db 2643 CCAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2584
QY 290 ATGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 346
Db 2583 ATGGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2524
QY 347 TGGATCTACAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 406
Db 2523 TTGACTTGACACCCCTCCATTTATCATTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2464
QY 407 GCTTGGGCAAGGAGTCTTCCCGAGGCTACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 466

Db	2463	GTCAATGGCAATCGAGTTTGGCCCCCATATGTTACAAGATCAATTTACAGTCGGGATGGCACA	TCGCT	2404
Qy	467	GGATGGCCTGGAAGACCGCTGGGCTCAGGAGGTGATCTCAGGCAATAGGAGACCC	TGAGG	526
Db	2403	GGATCTCTTGGCGGAACCGTCATGGAAACAGGTGCTGGATGGAATAAGTAACCCCTATG		2344
Qy	527	GAGTGGTGTGAAGGACCTTGGGGCCCCCATATGGTTGGCCGAGCTGGTTTCGCTTCA	ACCCCC	586
Db	2343	ACATTTTCTTAAGGACTTGGAGCGCCCATTTGATAGCAGATTTGTCCGTTTCATTCC	AG	2284
Qy	587	GGCTGACCGGGTCATGACGCTCTCTCTCGGGGTAGAGCTCTATGCTGCTCTGAGGG		646
Db	2283	TCACCGACCACTCCATGAATGTGTATGAGAGTGGAGCTTTACGGCTGTGTCTGGCT	AG	2224
Qy	647	ATGGACTCTGTCTTACACGCCCTGTGGGCGAGACAATGTAATTTATCTAGG	-----	701
Db	2223	ATGGCTTGGTGTCTTACATGCTCCAGCTGGCGAGCAGTTTGTACTCCTTGAGGTTCCA		2164
Qy	701	CCGTGACTCTAACGACTCCACTATATGACGCACATACCGTGGCGGAGCTGCATATGGGG		760
Db	2163	TCATTTATCTGAATGATTTGTCTTATGATGGAGCTGTGGATACAGCATCACAGAAGG	-	2104
Qy	761	GTCTGGCCAGCTGGCAGATGTGTGTGGGCTGGATGACTTTAGGAAGATCAGGAGC		820
Db	2104	--CTAGGCCAATTAGCCGATGTGTCTGGCCCTGGAGATTTTACCACGACCCATGAAT		2047
Qy	821	TGGGGTGTGGCCAGCTATGACTATGTGGGATGGAGCAACAGACTTCTCAGTGGCT		880
Db	2046	ACACGTGTGGCCGCTATGACTATGTGGCTGGCGAAGCAGAGTGCCACCAANTGGCT		1987
Qy	881	ATGTGGAGTGGATTTGAGTTTGACCGGCTGAGGGCCTCCAGGCTATGCAGTCCACT		940
Db	1986	ACATTGAGATCATGTTTGAATTTGACCGCATCAGSAATTTTCACTACCATGAAGTCCACT		1927
Qy	941	GTAAACAATATGCACAGCTGGGAGCCGCTCTGCCCTGGCGGGTGGATGTGCTTTCGGC		1000
Db	1926	GCAACAACATGTTTGTCTAAAGGTGTGAAGATCTTTAAGGAGGTACAGTGTACTTCGGCT		1867
Qy	1001	GTGGCCCTGCATGGCTCTGGGAGGGAGCCCATCGCCACAACCTTAGGGGCAACCTGG		1060
Db	1866	CTG---AAGCCATGAGTGGGAACCTAATGCCATTTCTTCCCTTGTCTCGATGACG		1810
Qy	1061	GGACCCACAGAGCCCGGCTGTCTCAGTGCCCTTTGGCGGCGGTGTGGCTCGTTCCTGC		1120
Db	1809	TCAACCCAGTGTCTGGTTTGTACGGTGGCTCTCCACCCAGGAATGSCCAGTGCCATCA		1750
Qy	1121	AGTGGCCTTCTCTTTGGGGGCGCTGGTTACTCTTTCAGCGAATCTCCTTTCATCTCTG		1180
Db	1749	AGTGTCAAATACCAATTTTGCAATACCTCGATGATGTTTCAGTCAGATCACTTCCCATC		1690
Qy	1181	ATGTGGTGAACAATTCCTCTCGGCACTGGGAGGACCTTCCCGCAGCCCTCGGTGGC		1240
Db	1689	ATGTGCAATGTAAACAACACTGTGAAGCCCTGCCACCTCTCTTA-----		1644
Qy	1241	CGCCTGGCCCACTCCCAACAACCTTCAGCAGCTTGGAGCTGGAGCCAGGAGCCAGCAGC		1300
Db	1644	-----TGSCACCAACCACTATGATC		1624
Qy	1301	CCGTGGCCAGAGCCGAGGGAGCCGACCGCCATCTCATCGGCTGCTTGGTGCCCATCA		1360
Db	1623	CAATGCTTAAAGTTGATGACAGCAACACACTCGGATCTCTGATTGGTGTGGTGCCATCA		1564
Qy	1361	TCCTGCTCTGCTCATCATGTCCTCATCTCTGGGCTGCACCTGGCGGAGGCTCC		1420
Db	1563	TCCTTTATCTCTTGGCCATTTGTCTCATCTCTGGAGGAGCTTCTGGCAGAAATGC		1504
Qy	1421	TCACAAGAGGCTGAACGAGGGGTGTTGGAAAGAGGAGCTGACGGTTTCACTCTCTCTGCC		1480
Db	1503	TGGAGAGGCTTCTCGAGGATCTCTGATGATGAATGACATGACATGACGCTTTCCCTG		1444
Qy	1481	GGGACATATCTCATCAACAACCGCCCAAGTCTCTAGAGAGGCCACCCCGTACCAAGAGC		1540
Db	1443	GTGATTTCTAGCATGTTTCAACAAT-----AACCGCTCT		1411

Qy	1541	CCGGCCTCGTGGGAATCGGCCCACTCCGCTCCCTGTGTGCCCAATGCGCTGTGCGTTGC	1500
Db	1410	CATCACTAGTAGAACAAGGTCCTCACTGACTTACGATCGCATCTTT-----	1363
Qy	1601	TGCTCTCCAATCCAGCCTACCGCCTCTTCTTGGCCACTTACGCCCGTCCCGCTCGAGGCC	1660
Db	1363	-----CCCCCTCGCC	1354
Qy	1661	CGGGCCCCCCCACACCCGCTGGGCCAAACCCACCAACACCCAGGCCCTACAGTGGGACT	1720
Db	1353	CTGACTACCAAGAGCCATCCAGCTGATACGAAAACTCCAGAAATTTGCTCCAGGGGAGG	1294
Qy	1721	ATATGAGCCTGAGAACCCAGCGGCCCGCTCTTGCCCGCCACCTCCCCAGACACGCTCC	1780
Db	1293	AGGAGTCAGGCTGCAGCGGTGTGTGAAGCCAGTCCAGCCAGTGGCCCTGAGGGGTGC	1234
Qy	1781	CCCATTTGCGAGGCTGACATTTACCTGCAAGGGGTCAACCGGGGCAACACCTATG	1840
Db	1233	CCCATTTGAGAGCTGACATAGTGAACCTTCCAAAGGATGACAGGAGGCCAACATACT	1174
Qy	1841	CTGTGCTGCACTGGCCCCCAGGGGCACTG---GGGATGGCCCCCCCCAGAGTGGAATTC	1897
Db	1173	CAGTGCTGCGGTACCATGGAACCTGCTCTCAGGAAAAAGATGTGGCTGTGGAGAGTTC	1114
Qy	1898	CTGATCTCGACTCGCTTCAAGGAGAAGCTTGGCGAGGCCAGTTGGGGAGGTGCACC	1957
Db	1113	CCAGGAACTCTTAACCTTTCAAGAGAAGCTGGGAGAAGGACATTTGGGGAGGTTCACT	1054
Qy	1958	TGTTGAGTGCACAGCCCTCAAGATCTGGTCACTTTGATTTCCCTCTTAATGTGCGTA	2017
Db	1053	TCTGTAAGTGGAGGAATGGAAAAATTCAGGAAAAAGATTTTGGCCCTAGATGTCACTG	994
Qy	2018	AGGACACCTTTGCTGTAGCTGTCAAGATCTTACGCCAGATGCCACCAAGAAATGCCA	2077
Db	993	CCAACAGACGCTGCTCTGGTGGCTGTGAAATGCTCCGAGCAGATGCCAACAAAGAAATGCCA	934
Qy	2078	GGAATGATTTCTTGAAGAGGTGAAGATCATGTCAGAGCTCAAGAGCCCAACCAATCATTC	2137
Db	933	GGAATGATTTCTTAAGAGATAAAGATCATGCTCGCTCAAGGACCCCAACATCATCC	874
Qy	2138	GGCTGTGGCGTGTGTGTGAGGAGCAGCCCTCTGATGATTAATGCTACATGACATGAGA	2197
Db	873	ATCTATTATCTGTGTATCACTGATGACCTCTCTGTATGATCACTAAATACATGAGA	814
Qy	2198	ACGGCACCCTAACCACTGTTCTCTAGTGGCCACCACTGGAGGACAAGCAGCCGAGGGGG	2257
Db	813	ATGGAGATCTCAATCAGTTTCTTCCCGCAGAGCCCCCTAATCTTCCTCCAGCG---	756
Qy	2258	CCCTTGGGGAGGGCAGGTGCGCAGGGGCCCACTCAGCTACCCAAATGCTGTGCATG	2317
Db	756	-----ATGTACGCACCTGTCAAGTACACCAATCTGAAGTTTA	721
Qy	2318	TGGCAGCCAGATCGCTCGGCATGCGTATCTGGCCACTCAACTTGTACATCGG	2377
Db	720	TGGTACCCAAATTTGCTCTGGCATGAAGTACTTTCCTCTCTTAATTTGTTCACCGAG	661
Qy	2378	ACCTGGCCAGCGAACTGCTAGTTGGGAAAAATTTACCATCAAAATCCGACACTTG	2437
Db	660	ATCTGGCCACACAACTGTTTAGTGGTGAAGAACTACACAAATCAAGATAGTCACTTTG	601
Qy	2438	GCATGACCGGAACCTCTATGCTGGGGACTATTACGTGTGCAGGGCCGGCAGTGTGTC	2497
Db	600	GAATGAGCAGGAACCTGTACAGTGGTGAATATTACCGGATCCAGGGCCGGCGAGTCTCC	541
Qy	2498	CCATCCCTGGATGGCTGGAGTGCATCCTCATGGGAAGTTTACAGTCTCGAGTACG	2557
Db	540	CTATCCCTGGATGCTTGGGAGATATCTTCTGGCAAGTTCACTACAGCAAGTGATG	481
Qy	2558	TGTGGGCTTTGTGTGAACCTGTGGGAGGTGCTGATGCTGTGTAAGGCCCAAGCCCTTTG	2617
Db	480	TGTGGGCTTTGGGGTACTTTTGTGGGAGACTTTTCACTTTGTCAAGAACAAGCCCTATT	421


```

; FILING DATE: 20-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/157563
; FILING DATE: 23-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 00, 000
; REFERENCE/DOCKET NUMBER: 854C1PIC2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2820 bases
; TYPE: nucleic acid
; STRANDEDNESS: Double
; TOPOLOGY: linear
US-08-441-104A-4

```

Db 2157 ATGGACCTGATGCCAA-----GCTGCTGGCTGGTGGGAGGATGTGGCTCCAG 2204
Qy 2288 CCACCATCAGTACCCAAATGCTGCTGATGTCAGCAGCCAGATCGCTCCGGCATCGCT 2347
Db 2205 GCCCCTGGGTGGGGCAGCTGTGGCCCTGCTGCTAGCCAGGTGCTGGGGGATGGTCT 2264
Qy 2348 ATCTGGCCACATCAACTTTGTACATCGGACCTGGCCAGCGGAACTGCTAGTTGGGG 2407
Db 2265 ACCTGGGGGTCTGCAATTTGTGACCGGGACCTGGCCACACCAACTGCTAGTGGGCC 2324
Qy 2408 AAAATTTCAACCATCAAAATCCAGACTTTTGGCATGAGCGGAACTTATGCTGGGGACT 2467
Db 2325 AGGACTGGTGGTCAAGATTGGTGAATTTGGCATGAGCAGGATATCTACAGCACCGAT 2384
Qy 2468 ATTACCGTGTGAGGGCGGGCAGTGTGCTCCATCGCTGGATGGCTGGGATGTCATCC 2527
Db 2385 ATTACCGTGTGGAGCGCGCACCATGCTGCCATTCGCTGATGCGCCCGAGAGCATCC 2444
Qy 2528 TCATGGGAAGTTTACGACACTTCCAGTGTGACGTGTGGGCTTTGTGTGACCTGTGGGAGG 2587
Db 2445 TGTACCTAAGTTTACACACGAGAGCGACGTGTGGAGCTTTCGGCGTGGTCTCTGGGAGA 2504
Qy 2588 TGTGTGATGCTGTGAGGGCGCGCCCTTTTGGCGAGCTCACCGACGAGAGGTATCGA 2645
Db 2505 TCTTCACTAGGCAAG---CAGCCCTGTGACCATCTCTCCAAACAGGAGGCAATCGA 2559

RESULT 11

US-08-306-691B-23
; Sequence 23, Application US/08306691B
; Patent No. 5734039

GENERAL INFORMATION:

; APPLICANT: Calabretta, Bruno
; APPLICANT: Skorski, Tomasz
; TITLE OF INVENTION: ANTISENSE
; TITLE OF INVENTION: OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seidel, Gonda, Lavorina & Monaco, P.C.
; STREET: Two Penn Center, Suite 1800
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19102

COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 720 kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1

CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/306,691B
; FILING DATE: September 15, 1994
; CLASSIFICATION: 514

PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 8321-8
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: NO. 5734039e

INFORMATION FOR SEQ ID NO: 23:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 2301 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear

US-08-306-691B-23

Query Match 5.3%; Score 193.6; DB 2; Length 2301;
Best Local Similarity 59.4%; Pred. No. 3.1e-34;
Matches 391; Conservative 0; Mismatches 249; Indels 18; Gaps 3;
Qy 1988 TCAGTCITGATTTCCCTTAAATGTGTAGGACACCTTTTCTGGTGTAGCTGTCAAGA 2047
Db 1275 TCCTTGTGTAGTGGACAACTCTCTGCTGAGCAGGACAGATGCTGTGTGTCTCAAG- 1334
Qy 2048 TCTTACGGCCAGATGCCACCAAGAAATCCAGGAATGCTTCTTCTTAAAGAGGTGAGATCA 2107
Db 1334 --GACTGAAGAGAGGCTCCAGAGATGCTCGCAGGACTTCCAACTGAGGCTGAGTGC 1391
Qy 2108 TGTGAGGCTCAAGAGACCCAAACATCATTCGGCTGCTGGCGTGTGTGTGAGCAGCAGC 2167
Db 1392 TCACCATGCTCAGCAGCAGCAGCATGCTGCGCTTCTTCGGCGTCTGCACCGAGGCGCC 1451
Qy 2168 CCTCTGCTGATTTACTGACTACATGAGAGACGGGACCTCAACCACTTCTCTAGTGGCC 2227
Db 1452 CCCTGCTCATGCTTCTGAGTATATGCGCACGGGAGCTCAACCGCTTCTCTCGATCCC 1511
Qy 2228 ACCAGCTGGAGGACAAAGCAGCAGCGAGGGGCCCTTGGGAGCGGACGCTGCGCAGGGCC 2287
Db 1512 ATGGACCCGATGCCAA-----GCTGCTGGTGTGGGAGGATGTGGCTCCAG 1559
Qy 2288 CCACCATCAGTACCCCAATGCTGCTGCTGTCAGCAGCCAGATCGCTCCGGCATGCGCT 2347
Db 1560 GCCCCTGGGTCTGGGCGAGCTGCTGGCCGTGGCTAGCCAGGTGCTGCGGGATGTGT 1619
Qy 2348 ATCTGGCCACACTCAACTTTTACATCGGACCTTGGCAGCGGAGTGTGTGTAGTGGGG 2407
Db 1620 ACCTGGCGGTGCTGCTATTTGTGACCGGGACCTTGGCAGCAGCACTGTCTAGTGGGCC 1679
Qy 2408 AAAATTTCAACCATCAAAATCCAGACTTTTGCATGAGCCGGAACCTCTATCTGGGACT 2467
Db 1680 AGGACTGGTGGTCAAGATTGGTATTTGGCATGAGCAGGATATCTACAGCAGCAGCT 1739
Qy 2468 ATTACCGTGTGAGGGCGGGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2527
Db 1740 ATTACCGTGTGAGGGCGGGCAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1799
Qy 2528 TCATGGGAAGTTTACGACTCGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2587
Db 1800 TGTACCGTAACTTCCACCGAGAGAGCAGCTGTGGAGCTTTCGGCGTGGTGTCTGGGAGA 1859
Qy 2588 TGTGTGATGCTGTGAGGGCCAGCCCTTGGGAGCTTCCAGCAGCAGCAGGATCAATCA 2645
Db 1860 TCTTCACTAGGCAAG---CAGCCCTGTGACCATCTCTCCAAACAGGAGGCAATCGA 1914

RESULT 12

PCT-US93-06251-78

; Sequence 78, Application PC/TUS9306251

GENERAL INFORMATION:

; APPLICANT: Wickstrom, Eric and Rife, Jason P.
; TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing
; TITLE OF INVENTION: Stereospecific Alkylphosphonates and Arylphosphonates
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: NY
; COUNTRY: USA
; ZIP: 11530

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/06251
; FILING DATE: 19930630
; CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:
NAME: Digilio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 8586
TELEPHONE: 516-742-4343
TELEFAX: 516-742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 78:
SEQUENCE CHARACTERISTICS:
LENGTH: 2301 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PCT-US93-06251-78

Query Match 5.3%; Score 193.6; DB 5; Length 2301;
Best Local Similarity 59.4%; Pred. No. 3.1e-34;
Matches 391; Conservative 0; Mismatches 249; Indels 18; Gaps 3;
QY 1988 TCAGTCTTGATTTCCCTTAAATGTCGCTAAGGACACACCTTTCTGCTAGCTGTCAAGA 2047
DB 1275 TCCTTGTGCTGAGTGCCACAACTCTCGCTGAGCAGGACAGATGCTGGTGGCTGTCAAG- 1334
QY 2048 TCTTACGGCCAGATGCCACCAAGAAATGCCAGGAATGATTTCTGAAAGAGGTGAAGATCA 2107
DB 1334 --GCACTGAAGAGGCGTCCGAGAGTGTCCGCGAGGACTTCAACGTGAGGCTGAGTGC 1391
QY 2108 TGTGAGGCTCAAGGACCCCAACATCATTCGGCTGCTGGGGTGTGTGTGAGGACGACC 2167
DB 1392 TCACCATGCTGCAGCAGCAGCAGCATCTGCTGCTTCTTGGGGTGTGCACCGAGGGCGCC 1451
QY 2168 CCTCTGCTGATTTACTACTACTATGAGAGACGGGACCTCAACAGTTCTCTAGTGGCC 2227
DB 1452 CCTCTGCTGATGCTTCTGAGTATATGCGGACGCGGGGACCTCAACCGTTCCTCCGATCCC 1511
QY 2228 ACCAGCTGGAGGACAGGACGCGAGGGGGCCCTGGGGAGCGGAGCTGCGCAGGGGC 2287
DB 1512 ATGGACCGATGCCAA-----GCTGCTGGCTGGTGGGAGGATGGCTCCAG 1559
QY 2288 CCACCATCAGCTACCAATGCTGTGATGTGGCAGCCCGGATCGCCTCCGGCATCGCT 2347
DB 1560 GCCCCTGGTCTGGGGCAGCTGTGGCGGTGGTAGCCAGTCTGCTGGGGATGGTGT 1619
QY 2348 ATCTGGCCACTCAACTTTGTACATCGGAGCTGGCCACCGGAACTGCCCTAGTTGGG 2407
DB 1620 ACCTGGCGGGTCTGCATTTTGTGACCGGGACCTGGCCACAGCAACTGTCTAGTGGGC 1679
QY 2408 AAAATTTCCACATCAAAATCGCAGACTTTGGCATGAGCCGGAACCTCTATGCTGGGACT 2467
DB 1680 AGGAGTGGTCAAGATTTGGTATTTGGCATGAGCAGGAGATATCTACAGCAGCACT 1739
QY 2468 ATTACCGTGTGACGGCGGCGCAGTGTGCCATCCGCTGGATGCCCTGGGAGTGCATCC 2527
DB 1740 ATTACCGTGTGGAGGCGCGCACCATGCTGCCATTCGCTGGATGCCGCCGAGAGCATCC 1799
QY 2528 TCATGGGAAAGTTACAGACTCGGAGTACGCTGTGGGCCCTTTGGTGTGACCTGTGGAGG 2587
DB 1800 TGATCCGTAAGTTACACCGAGAGACGCTGTGGAGTCTCGGCTGGTGTCTGGGAGA 1859
QY 2588 TGCTGATGCTGTGAGGCCAGCCCTTTGGGAGCTCACCGACGAGCAGGTCTATCGA 2645
DB 1860 TCTTCACTACGGCAAG---CAGCCCTGGTACCAGCTCTCCAAACAGGAGCAATCGA 1914

RESULT 13
US-08-271-454-1
; Sequence 1, Application US/08271454
; Patent No. 5601820
; GENERAL INFORMATION:
; APPLICANT: Brodeur, Garrett M.

APPLICANT: Nakagawara, Akira
TITLE OF INVENTION: Compositions and Methods of Making and
; TITLE OF INVENTION: Using Human Full Length TRK-B
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 5601820rls
STREET: One Liberty Place, 46th floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/271,454
FILING DATE: 07-JUL-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: DeLuca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: CH-0535
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3707 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: both
MOLECULE TYPE: CDNA
US-08-271-454-1

Query Match 5.2%; Score 190; DB 1; Length 3707;
Best Local Similarity 60.7%; Pred. No. 2.3e-33;
Matches 357; Conservative 0; Mismatches 210; Indels 21; Gaps 2;

QY 2059 GATGCCACCAAGATGCCAGGAATGATTTCTGAAAGAGTGAAGATCATGTCGAGGCTC 2118
DB 2077 GATGCCAGTGAATGACAGCAAGGACTTCCACCGTGGGGGAGCTCTGTACCAACCTC 2136
QY 2119 AAGGACCCAAACATCATTCGGCTGCTGGCGGTGTGTGTCAGGACGACCCCTCTGCATG 2178
DB 2137 CAGCATGACATCATGTCAGTTCTATGGCGTCTGCGTGGAGGGGCGCCCTCATCATG 2196
QY 2179 ATTACTGATGATGAGAACCGGACCTCAACAGTTCTCTCAGTCCGACAGCTGGAG 2238
DB 2197 GTCTTTGATGATGAGCATGGGACCTCAACAAGTTCTCAGGGCAGACGGCCCTGAT 2256
QY 2239 GACAGGACGCGAGGGGCGCCCTGGGGACGGGAGGTGCGCAGGGGCGCCACCATCAGC 2298
DB 2257 GCGTCTGATGGTGGGCAACCGCCGCGGAACTGACGCACTGCG----- 2306
QY 2299 TACCCAACTGCTGCTGATGTCGACGCGGAACTGCTAGTTGGGAAATTTTACC 2358
DB 2306 -----AGATGCTGATATAGCCAGAGATCGCGCGGGCATGGTTACCTGGCTCC 2358
QY 2359 CTCACATTTGATGTCGACGCGGAACTGCTAGTTGGGAAATTTTACC 2418
DB 2359 CAGCACTTGTGACCGGATTTGGCCACGAGAACTGCTGCTGGGAGAACTTGTG 2418
QY 2419 ATCAAAATCGAGACTTTGGCATGAGCGGAACTCTTATGTGGGAACTTATACCGTGTG 2478
DB 2419 GTGAAATCGGGAGCTTTGGGATGTCCCGGAGCTGTACAGCACTGACTTACAGGGTC 2478
QY 2479 CAGGGCGGGAGCTGCTGCCATCGCTGGATGGCTGGGAGTGTCTCATGCGGAG 2538
DB 2479 GGTGGCCACACAACTGCTGCCATTCGCTGGATGGCTCCAGAGAGCATCATGTACAGGAA 2538

[illegible]

RESULT 14
US-08-286-305A-6
; Sequence 6, Application US/08286305A
; Patent No. 5766863
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J.
; APPLICANT: Mark, Melanie R.
; APPLICANT: Sadick, Michael D.
; APPLICANT: Shelton, David L.
; APPLICANT: Wong, Wai Lee Tan

	Query Match	5.2%	Score 190;	DB 3;	Length 3060;
	Best Local Similarity	60.7%	Pred. No. 2.2e-33;		
	Matches 357;	Conservative 0;	Mismatches 210;	Indels 21;	Gaps 2;
QY	2059	GATGCCACCAAGATGCCAGGAATGATTCTCTGAACAGGTGAAGATCATGTCTGAGGCTC	2118		
Db	2107	GATGCCAGTGAATATGCACGAAGAGACTTCCACGCTGAGCCGAGCTCTCTGACCAACCTC	2166		
QY	2119	AAGGACCCAAACATCATTCGGCTCTCTGGCGTGCTGTGCAGGACGACCCCTCTCGATG	2178		
Db	2167	CAGCATGACCATCGTCAAGTTCTTATGGCGTCTGCGTGGAGGCGACCCCTCTCATCTG	2226		
QY	2179	ATTACTGACTACATGGAGAACGGGCGACTCAACCAGTTCTCTAGTGCACCAAGCTGGAG	2238		

Db	2227..GTCTTTTGAGTATACAAAGCATGGGAGCCTCAACAAGTTCTCTACAGGCACACAGGCCCTGAT	2285
QY	2239 GACAAGCGACGAGGGGCCCTTGGGACGGGAGGCTGCGCAGGGGCCACCATCAGC	2298
Db	2287 GCGTGCTGATGGCTGAGGGCAACCCGCCACGGAAGTGACGAGTCGC-----	2336
QY	2299 TACCCAAATGCTGCTGCATGTGGCAGCCACAGATCGCCTCCGSCATGCGCTATCTGGCCACA	2358
Db	2336 -----AGATGCTGCATATAGCCACAGCATCGCCGGGCGATGGTCTACCTGGCGTCC	2388
QY	2359 CTCAACTTTGTACATCGGGAGCTGGCCACGCGGAACCTGCCTAGTTGGGGAATTTACCC	2418
Db	2389 CAGCACTTCGTGCACGGGATTTGGCCACAGAACTGCCTGTGTCGGGAGAACTTGCTG	2448
QY	2419 ATCAAATCGCAGACTTTGGCATGACCGGAACCTCTATGTCGGGGAATTTACCGTGTG	2478
Db	2449 GTGAAATCGGGACTTTGGGATGTCCCGGACGTGTACAGCACTGACTACTACGGTTC	2508
QY	2479 CAGGGCCGGCAGTGCTGCCATCGCTGGATGGCCCTGGGAGTGCACTCTCATGGGGAAG	2538
Db	2509 GGTGGCCACACAATGCTGCCATTCGCTGGATGCTCCAGAGASCATCATGTACAGGAAA	2568
QY	2539 TTACAGACTGCGAGTGACGTGTGGGCCCTTTGGTGACCCCTGTGGGAGGTGCTGATGCTC	2598
Db	2569 TTCACGACGGAAAGCAGCTGTGGAGCCTGGGGTCTGTTGTGGGAGATTTTCA---CC	2625
QY	2599 TGTAGGGCCCCAGCCCTTTGGCAGCTCACCGACGAGCAGGTTCATCGAG	2646
Db	2626 TATGGCAACAGCCCTGTTACCAGCTGTCAACAAATAGAGTGTATAG	2673

RESULT 15

US-08-359-705B-1
; Sequence 1, Application US/08359705B
; Patent No. 5844092
; GENERAL INFORMATION:
; APPLICANT: Presta, Leonard G.
; APPLICANT: Shelton, David L.
; APPLICANT: Urrer, Roman
; TITLE OF INVENTION: Human trk Receptors and Neurotrophic Factor Inhibitors
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/359.705B
; FILING DATE: 20-Dec-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/286846
; FILING DATE: 08/10/94
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/215139
; FILING DATE: 03/18/94
; ATTORNEY/AGENT INFORMATION:
; NAME: Torchia, Phd., Timothy E.
; REGISTRATION NUMBER: 36,700
; REFERENCE/DOCKET NUMBER: P0873P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-8674
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:

Search completed: November 4, 1999, 07:29:20
Job time: 6308 sec

Db 496 CCCGCCACAGAGGTTGGAGAGCAGTACCGGGGATGGGGCTTGGTGGCCCGCAGGCTCGG 555
Qy 317 TGTGTTCCCAAGAGAGGAGTACTTGCAGGTGGATCTACAAAGCTGCACTGCTGGTCTC 376
Db 556 TGTGTTCCCAAGAGAGGAGTACTTGCAGGTGGATCTACAAAGCTGCACTGCTGGTCTC 615
Qy 377 TGGTGGGACCCAGGAGCGCATCCGGGGGCTGGCAAGAGTCTTCCCGAGCTACC 436
Db 616 TGGTGGGACCCAGGAGCGCATCCGGGGGCTGGCAAGAGTCTTCCCGAGCTACC 675
Qy 437 GCGTGGCTTACTCCCGGATGGTCCCGCTGGATGGCTGGAAGGACCGCTGGGGTCAGG 496
Db 676 GCGTGGCTTACTCCCGGATGGTCCCGCTGGATGGCTGGAAGGACCGCTGGGGTCAGG 735
Qy 497 AGGTGATCTCAGGCAATGAGGACCTGAGGAGTGGTCTGAAAGGACCTTGGGGCCCCCA 556
Db 736 AGGTGATCTCAGGCAATGAGGACCTGAGGAGTGGTCTGAAAGGACCTTGGGGCCCCCA 795
Qy 557 TGGTGGCCGACTGGTTCGCTTACCCCGGGCTGACCGGGTCAAGAGCTTGTCTGTCG 616
Db 796 TGGTGGCCGACTGGTTCGCTTACCCCGGGCTGACCGGGTCAAGAGTGTCTGTCGTCG 855
Qy 617 GGTAGAGCTATGCTGCTCTGGAGGATGGACTCTCTTACACCGCCCTGTGG 676
Db 856 GGTAGAGCTATGCTGCTCTGGAGGATGGACTCTCTTACACCGCCCTGTGG 915
Qy 677 GGCAGACATGTTATCTGAGCGCTGTACCTCAAGACTCCACCTATGAGGACATA 736
Db 916 GGCAGACATGTTATCTGAGCGCTGTACCTCAAGACTCCACCTATGAGGACATA 975
Qy 737 CCGTGGCGGACTGCAGTATGGGGTCTGGGCCAGCTGGCAGATGGTGGGGCTGG 796
Db 976 CCGTGGCGGACTGCAGTATGGGGTCTGGGCCAGCTGGCAGATGGTGGGGCTGG 1035
Qy 797 ATGACTTTAGGAAGTACAGAGCTGGGGTCTGGCCAGGCTATGCTATGTTGGATGGA 856
Db 1036 ATGACTTTAGGAAGTACAGAGCTGGGGTCTGGCCAGGCTATGCTATGTTGGATGGA 1095
Qy 857 GCAACACAGCTTCTCCAGTGGCTATGTTGAGATGGAGTTGAGTTTACCGGCTGAGG 916
Db 1096 GCAACACAGCTTCTCCAGTGGCTATGTTGAGATGGAGTTGAGTTTACCGGCTGAGG 1155
Qy 917 CTTCCAGGCTATGAGTCCACTGTAAACACATGCACAGCTGGGAGCCCGTCTGCTG 976
Db 1156 CTTCCAGGCTATGAGTCCACTGTAAACACATGCACAGCTGGGAGCCCGTCTGCTG 1215
Qy 977 GCGGGGTGGAATGCTCTCCGGGTGGCCCTGCCATGGCTGGGAGGGGAGCCCATGC 1036
Db 1216 GCGGGGTGGAATGCTCTCCGGGTGGCCCTGCCATGGCTGGGAGGGGAGCCCATGC 1275
Qy 1037 GCAACACTAGGGGCAACCTGGGGACCCAGAGCCCGGGCTGTCTCAGTGGCCCTTG 1096
Db 1276 GCCAAGCTAGGGGCAACCTGGGGACCCAGAGCCCGGGCTGTCTCAGTGGCCCTTG 1335
Qy 1097 GCGGCGGTGGCTGCTTCTGAGTGGCCGCTTCTCTTTCGGGGCCCTGGTTACTCT 1156
Db 1336 GCGGCGGTGGCTGCTTCTGAGTGGCCGCTTCTCTTTCGGGGCCCTGGTTACTCT 1395
Qy 1157 TCAGGGAATCTCTCTATCTGATGTGGTGAACAATCTCTCCGGCAGCTGGAGGCA 1216
Db 1396 TCAGGGAATCTCTCTATCTGATGTGGTGAACAATCTCTCCGGCAGCTGGAGGCA 1455
Qy 1217 CTTCCCGCCAGCCCTGGTGGCGCTGGCCAGCTCCCACTCCCACTTACAGAGCTGG 1276
Db 1456 CTTCCCGCCAGCCCTGGTGGCGCTGGCCAGCTCCCACTCCCACTTACAGAGCTGG 1515
Qy 1277 AGCTGGAGCCAGAGCCAGAGCCCGTGGCCAAAGCCGAGGAGCCCGACCCCATCC 1336
Db 1516 AGCTGGAGCCAGAGCCAGAGCCCGTGGCCAAAGCCGAGGAGCCCGACCCCATCC 1575
Qy 1337 TCATGGCTGCTGGTGGCCATCTCTGCTGCTCTCATCTATCTGCTGCTCT 1396
Db 1576 TCATGGCTGCTGGTGGCCATCTCTGCTGCTCTCATCTATCTGCTGCTCT 1635

Qy 1397 GCGGCTGCTGCTGCGCAGGCTCTCAGCAAGCTGAAAGGAGGTGTGGAAGAGGAGC 1456
Db 1636 GCGGCTGCTGCTGCGCAGGCTCTCAGCAAGCTGAAAGGAGGTGTGGAAGAGGAGC 1695
Qy 1457 TGACGGTTCACCTCTCTGCTGCGTGGGACATATCTCATCAACAGCGCCAGGTCTTA 1516
Db 1696 TGACGGTTCACCTCTCTGCTGCGTGGGACATATCTCATCAACAGCGCCAGGTCTTA 1755
Qy 1517 GAGAGCCACCCCGTACAGGAGCCCGGCTCTGTTGGAATCCCGCCACTCCGCTCCCT 1576
Db 1756 GAGAGCCACCCCGTACAGGAGCCCGGCTCTGTTGGAATCCCGCCACTCCGCTCCCT 1815
Qy 1577 GTGTGCTCCCAATGGCTCTGCTGCTCTCCAAATCCAGCCTACCGCTCTCTTGGCCA 1636
Db 1816 GTGTGCTCCCAATGGCTCTGCTGCTCTCCAAATCCAGCCTACCGCTCTCTTGGCCA 1875
Qy 1637 CTTAGCGCGCTCCCGCTCGAGGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1696
Db 1876 CTTAGCGCGCTCCCGCTCGAGGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1935
Qy 1697 ACACCCAGGCTACAGTGGGACTATATGAGCCTGAGAGCCAGGCGCCGCTTCTGCG 1756
Db 1936 ACACCCAGGCTACAGTGGGACTATATGAGCCTGAGAGCCAGGCGCCGCTTCTGCG 1995
Qy 1757 CCCCACCTCCCGCAGAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1816
Db 1996 CCCCACCTCCCGCAGAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2055
Qy 1817 GCGTACCGGGGGCAACACCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1876
Db 2056 GCGTACCGGGGGCAACACCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2115
Qy 1877 GCGCCCGCAGAGTGGATTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1936
Db 2116 GCGCCCGCAGAGTGGATTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2175
Qy 1937 GCGAGTTTGGGAGGTGACCTGCTGAGGCTGAGAGCCCTCAAGATCTGCTGCTGCTGCTG 1996
Db 2176 GCGAGTTTGGGAGGTGACCTGCTGAGGCTGAGAGCCCTCAAGATCTGCTGCTGCTGCTG 2235
Qy 1997 ATTTCCTTAAATGCTGCTGAGGACACCTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTG 2056
Db 2236 ATTTCCTTAAATGCTGCTGAGGACACCTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTG 2295
Qy 2057 CAGATGCCACCAAGATGCGAGCTTCTCTGTTCTCCAGGAATGATTTCTGAAAGAGG 2098
Db 2296 CAGATGCCACCAAGATGCGAGCTTCTCTGTTCTCCAGGAATGATTTCTGAAAGAGG 2355
Qy 2099 TGAAGATCATGTCGAGGCTCAAGGACCCCAACATCATTCGGCTGCTGGCGTGTGTGTC 2158
Db 2356 TGAAGATCATGTCGAGGCTCAAGGACCCCAACATCATTCGGCTGCTGGCGTGTGTGTC 2415
Qy 2159 AGGAGACCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2218
Db 2416 AGGAGACCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2475
Qy 2219 TCAGTGGCCACAGCTGGAGGACAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 2278
Db 2476 TCAGTGGCCACAGCTGGAGGACAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2535
Qy 2279 GCGAGGGGCGCCACCATCAGCTACCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2338
Db 2536 GCGAGGGGCGCCACCATCAGCTACCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2595
Qy 2339 GCATGGCTATCTGGCAGACTCACTTGTACATCGGAGCTGGGAGCTGGGAGCTGGGAGCTGG 2398
Db 2596 GCATGGCTATCTGGCAGACTCACTTGTACATCGGAGCTGGGAGCTGGGAGCTGGGAGCTGG 2655
Qy 2399 TAGTTGGGAAAAATTTACCATCAAAATCGCAGACTTTGGCATGAGCGGAACTCTATG 2458
Db 2656 TAGTTGGGAAAAATTTACCATCAAAATCGCAGACTTTGGCATGAGCGGAACTCTATG 2715

QY 2459 CTGGGAGCTATTACGCTGTGAGGCGCGGCGAGTGTGCGCCATCCGCTGGATGCCCTGGG 2518
DB 2716 CTGGGAGCTATTACGCTGTGAGGCGCGGCGAGTGTGCGCCATCCGCTGGATGCCCTGGG 2775
QY 2519 AGTGCATCTCATGGGAAAGTTACAGACTGCGAGTGTGAGTGTGCGCTTTGGTGTGACCC 2578
DB 2776 AGTGCATCTCATGGGAAAGTTACAGACTGCGAGTGTGAGTGTGCGCTTTGGTGTGACCC 2835
QY 2579 TGTGGAGGTGCTGATGCTGTGTAGGCGCCAGCCCTTTGGGAGCTCATCCGACGAGCAGG 2638
DB 2836 TGTGGAGGTGCTGATGCTGTGTAGGCGCCAGCCCTTTGGGAGCTCATCCGACGAGCAGG 2895
QY 2639 TCATCGAAGCGGGGGAGTCTTCGCGGACACAGGCGCGGCGAGTGTACCTGTCCCGGC 2698
DB 2896 TCATCGAAGCGGGGGAGTCTTCGCGGACACAGGCGCGGCGAGTGTACCTGTCCCGGC 2955
QY 2699 CGCCTGCTGCCCGGAGGCTATATAGCTGTATGCTTCGCTGCTGAGCGGGAGTCTG 2758
DB 2956 CGCCTGCTGCCCGGAGGCTATATAGCTGTATGCTTCGCTGCTGAGCGGGAGTCTG 3015
QY 2759 AGCAGCGACCACTCTTTCCAGCTGCTATGCTGCTGCTGCGAGAGTGCACCTCAACACGG 2818
DB 3016 AGCAGCGACCACTCTTTCCAGCTGCTATGCTGCTGCTGCGAGAGTGCACCTCAACACGG 3075
QY 2819 TGTGAATCACACATCCAGCTGCCCTCCCTCAGGAGTGTATCCAGGGAAGCCAGTGCACA 2878
DB 3076 TGTGAATCACACATCCAGCTGCCCTCCCTCAGGAGTGTATCCAGGGAAGCCAGTGCACA 3135
QY 2879 CTAAACAGAGGAGGACAAATGCGACCTCTGCTGCCCTCCCTCCGACAGCCCATCCTCT 2938
DB 3136 CTAAACAGAGGAGGACAAATGCGACCTCTGCTGCCCTCCCTCCGACAGCCCATCCTCT 3195
QY 2939 AATAGAGCAGTGTGAGCTGC----- 2959
DB 3196 AATAGAGCAGTGTGAGCTGCAGGTGGGCTGGGCGCCACCCAGGAGCTGATGCCCTTCTC 3255
QY 2959 ----- -AGAGCCCGCTGTGC 2972
DB 3256 CCCTCTCTGGACACACTCTATGCTGCCCTCTCTGTTCTTCTTCTTCTTCTTCTTCTG 3315
QY 2973 CCCACCCAGCTGGTCTGTGATGGATGCTTCCTCCAGCCCTCTTCTAGCCATCTCTGGG 3032
DB 3316 CCCACCCAGCTGGTCTGTGATGGATGCTTCCTCCAGCCCTCTTCTAGCCATCTCTGGG 3075
QY 3033 AAGGTGGGAGAAATATAGATAGACACTGGACATGGCCCATTTGGAGCACTTGGGCCCC 3092
DB 3376 AAGGTGGGAGAAATATAGATAGACACTGGACATGGCCCATTTGGAGCACTTGGGCCCC 3435
QY 3093 ACTGACAACTGATCTCTGGAGAGTGTGCTGCGCCCCAGCTTCTCTCCCTGTGCAC 3152
DB 3436 ACTGACAACTGATCTCTGGAGAGTGTGCTGCGCCCCAGCTTCTCTCCCTGTGCAC 3494
QY 3153 ACATGGACCCCTGCTGAGTCTGGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3212
DB 3495 ACATGGACCCCTGCTGAGTCTGGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3554
QY 3213 TTTCTTGTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3272
DB 3555 TTTCTTGTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3614
QY 3273 GAACACTGGACCTGGGGGTAGCGCGCCCGCCAGCCCTCAGTACCCCACTTCCCACTG 3332
DB 3615 GAACACTGGACCTGGGGGTAGCGCGCCCGCCAGCCCTCAGTACCCCACTTCCCACTG 3674
QY 3333 CAGTCTTCTAGCTAGACTTCTTCTAGCCTATAGTCTTCTGAGTAAATATATGGGATT 3392
DB 3675 CAGTCTTCTAGCTAGACTTCTTCTAGCCTATAGTCTTCTGAGTAAATATATGGGATT 3734
QY 3393 GGGGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3452
DB 3735 GGGGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3794
QY 3453 CACATTGATTTTCTATTAATCACTTGGGTTTGTACATTTTGGGGGGGAGAGACACAGAT 3512

DB 3795 CACATTGATTTTCTATTAATCACTTGGGTTTGTACATTTTGGGGGAGAGACACAGAT 3854
QY 3513 TTTTACACTAATATATGAGCTAGCTTGGGCAATTTAAATCCCTGCTAGGAGGAGTGA 3572
DB 3855 TTTTACACTAATATATGAGCTAGCTTGGGCAATTTAAATCCCTGCTAGGAGGAGTGA 3914
QY 3573 ATAATAAGGTTGAGTTTCCACAAAAAATAAAAAA 3611
DB 3915 ATAATAAGGTTGAGTTTCCACAAAAAATAAAAAA 3953
RESULT 2
QY2520 ID Q92520 standard; cDNA to mRNA; 3962 BP.
AC Q92520:
DE 26-NOV-1995 (first entry)
KW Human mammary carcinoma kinase 10 (MCK-10) cDNA.
KW Mammary carcinoma kinase 10; transmembrane receptor;
KW receptor tyrosine kinase; cancer; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT 321..3080 /*tag= a
FT misc_difference 2315 /*tag= b
FT /*note= "some clones have 6 AA deletion here"
PN WO9514088-A.
PD 26-MAY-1995.
PF 16-NOV-1994; E03797.
PR 16-NOV-1993; US-153397.
PA (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
PI Alves FHE, Ullrich A;
DR WPI: 95-224054/29.
DR P-PSDB; R75502.
PT New nucleic acid encoding MCK-10 receptor tyrosine kinase - and
PT derived vectors, transformed cells, proteins and antibodies useful
PT for diagnosis and treatment of proliferative disease, esp. cancer,
PT and for screening modulators
PS Claim 4; Page 50-52; 115pp; English.
CC cDNA prep'd. from human breast cancer cell line MCF7 (ATCC HTB22) and
CC used in a PCR with two degenerate oligo primer pools based on
CC conserved sequences of the kinase domain of receptor tyrosine
CC kinases. One clone, designated MCK-10, was identified as novel RTK.
CC The PCR fragment was used to screen a lambda gt11 library of human
CC fetal brain cDNA. Several overlapping clones were identified. The
CC composite of these cDNA clones is given in Q92520 and the deduced AA
CC sequence in R75502. Some of the clones had a deletion of 6 AAs at
CC position 2315 in the MCK-10 sequence. MCK-10 has all the
CC characteristics of a receptor PTK (see R75502 FT). Screening of
CC human placental library yielded two cDNA clones MCK-10-1 and
CC MCK-10-2. One of the clones isolated from the human fetal brain
CC library cont'd. an additional 18 nts in the TK domain. The MCK-10 splice
CC isoforms have been designated MCK-10-1 (with an additional 111 bp between
CC nts 1832 and 1943); MCK-10-2 (without any insertions); MCK-10-3 (with the
CC additional 111 bp and 18 bp in the TK domain); and MCK-10-4 (with the
CC additional 18 bp). The predicted mol. wts. of MCK-10-1 and MCK-10-2
CC proteolipids are 101.13 and 97.17 kD respectively, and can thus be
CC subdivided into a 34.31 kD alpha subunit and 66.84 or 62.88 kD
CC beta subunits that contain the TK homology and alternative splice sites.
CC Sequence 3962 BP; 735 A; 1235 C; 1181 G; 811 T;
SQ

Query Match 94.8%; Score 3449.4; DB 1; Length 3962;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 3588; Conservative 0; Mismatches 6; Indels 105; Gaps 3;
QY 17 GTTGACTTGAAGGAATGCCAAGAGATGCTGCCCCACCCCTTAGGCCCGAGGATCAG 76
DB 256 GTTGACTTGAAGGAATGCCAAGAGATGCTGCCCCACCCCTTAGGCCCGAGGATCAG 315
QY 77 GAGCTATGGGACGAGGCGCTGTCTTACTGCTGCTCTTGTGGTGGCAAGTGAG 136

Db 316 GAGCTATGGACAGAGCCCTGTCTATCTTTACTGCTGTCTCTTGGTGGCAAGTGGAG 375
QY 137 ATGCTGACATGAAGGACATTTTGATTCCTGCAAGTGGCCGCTATGCCCTGGGCATGCAAG 196
Db 376 ATGCTGACATGAAGGACATTTTGATTCCTGCAAGTGGCCGCTATGCCCTGGGCATGCAAG 435
QY 197 ACCGGACCATCCAGACATGACATCTCTGCTTCAGCTTCCTGGTCAGATTCACATGGCG 256
Db 436 ACCGGACCATCCAGACATGACATCTCTGCTTCAGCTTCCTGGTCAGATTCACATGGCG 495
QY 257 CCGCCACACAGCAGTTGGAGACAGTACGGGATGGGGCTGGTCCCGCAGAGGTGG 316
Db 496 CCGCCACACAGCAGTTGGAGACAGTACGGGATGGGGCTGGTCCCGCAGAGGTGG 555
QY 317 TGTTCCTCCAGAGGAGGAGTACTTGCAGTGGATCTACAAGACTGACACCTGGTGGCTC 376
Db 556 TGTTCCTCCAGAGGAGGAGTACTTGCAGTGGATCTACAAGACTTCCACCTGGTGGCTC 615
QY 377 TGTGGGCACCCAGGACAGGCATGCCGGGGGCTGGGCAAGGAGTTCTCCCGAGCTACC 436
Db 616 TGTGGGCACCCAGGACAGGCATGCCGGGGGCTGGGCAAGGAGTTCTCCCGAGCTACC 675
QY 437 GGCCTGGTTACTCCCGGATGGTCGGCTGGATGGCTGGAAGGACCGCTGGGGTCAAG 496
Db 676 GGCCTGGTTACTCCCGGATGGTCGGCTGGATGGCTGGAAGGACCGCTGGGGTCAAG 735
QY 497 AGGTGATCTCAGCAATAGAGACCTTGAGGAGTGGTGTGAAGACCTTGGGCCCCCA 556
Db 736 AGGTGATCTCAGCAATAGAGACCTTGAGGAGTGGTGTGAAGACCTTGGGCCCCCA 795
QY 557 TGTTCGCGACTGGTTCTGCTTACCCCGGGCTGACCGGGTCAAGCGGTCTGTCTGTC 616
Db 796 TGTTCGCGACTGGTTCTGCTTACCCCGGGCTGACCGGGTCAAGCGGTCTGTCTGTC 855
QY 617 GGTAGAGCTCTATGCTGCTCTGAGGAGTGGACTCTCTTACACCGCCCTGTGG 676
Db 856 GGTAGAGCTCTATGCTGCTCTGAGGAGTGGACTCTCTTACACCGCCCTGTGG 915
QY 677 GGCAGACAATGTATTTATCTGAGGCGGTACCTCAACGACTCCACTATGACGACATA 736
Db 916 GGCAGACAATGTATTTATCTGAGGCGGTACCTCAACGACTCCACTATGACGACATA 975
QY 737 CCGTGGCGGACTGCAATGAGGAGTGGGGTCTGGCCAGCTGGCAGTGGTGGGGCTGG 796
Db 976 CCGTGGCGGACTGCAATGAGGAGTGGGGTCTGGCCAGCTGGCAGTGGTGGGGCTGG 1035
QY 797 ATGACTTTAGGAAGTCAAGGAGTCCGGCTCTGGCCAGGCTATGACTATGGGATGA 856
Db 1036 ATGACTTTAGGAAGTCAAGGAGTCCGGCTCTGGCCAGGCTATGACTATGGGATGA 1095
QY 857 GCAACACAGCTTCTCCAGTGGCTATGTGGAGATGGAGTTGAGTTGACCGGCTGAGGG 916
Db 1096 GCAACACAGCTTCTCCAGTGGCTATGTGGAGATGGAGTTGAGTTGACCGGCTGAGGG 1155
QY 917 CTTTCAGGCTATGCAAGTCCACTGTAAACAATGACACAGCTGGAGCCCTCTGCCTG 976
Db 1156 CTTTCAGGCTATGCAAGTCCACTGTAAACAATGACACAGCTGGAGCCCTCTGCCTG 1215
QY 977 GCGGGTGGAAATCTCGCTTCCGGCTGGCCCTGCCATGGCTGGAGGGGAGCCCATGTC 1036
Db 1216 GCGGGTGGAAATCTCGCTTCCGGCTGGCCCTGCCATGGCTGGAGGGGAGCCCATGTC 1275
QY 1037 GCCAACACCTTAGGGGCAACCTTGGGGACCCAGAGCCCGGGCTGTCTCAGTGGCCCTTG 1096
Db 1276 GCCAACACCTTAGGGGCAACCTTGGGGACCCAGAGCCCGGGCTGTCTCAGTGGCCCTTG 1335
QY 1097 GCGGCGGTGGCTCGCTTCTGCACTGGCGCTTCTCTTTGGGGGGCCCTGGTACTCT 1156
Db 1336 GCGGCGGTGGCTCGCTTCTGCACTGGCGCTTCTCTTTGGGGGGCCCTGGTACTCT 1395
QY 1157 TCAGCGAAATCTCTTCACTCTGTATGTGGTGAACAATTCCTCTCGGACCTGGGAGGCA 1216
Db 1396 TCAGCGAAATCTCTTCACTCTGTATGTGGTGAACAATTCCTCTCGGACCTGGGAGGCA 1455

QY 1217 CTTTCGCGCAGAGCCCTGTGTGGCGGCTTGGCCCACTTCCCAACAACTTCAGCAGCTTGG 1276
Db 1456 CTTTCGCGCAGAGCCCTGTGTGGCGGCTTGGCCCACTTCCCAACAACTTCAGCAGCTTGG 1515
QY 1277 AGCTGGAGCCAGAGCCAGCAGCCGTCGCCAAGCCCGAGGGGAGCCCGCCATCC 1336
Db 1516 AGCTGGAGCCAGAGCCAGCAGCCGTCGCCAAGCCCGAGGGGAGCCCGCCATCC 1575
QY 1337 TCATCGGCTGCCCTGGTGGCCATCATCTGCTGTGTGTCTCATCTTGGCCTCATCTCT 1396
Db 1576 TCATCGGCTGCCCTGGTGGCCATCATCTGCTGTGTGTCTCATCTTGGCCTCATCTCT 1635
QY 1397 GCGGCTGTCATGCGCAGGCTCCTCAGCAAGGCTGAAAGGAGGGTGTGGAAAGAGGAGC 1456
Db 1636 GCGGCTGTCATGCGCAGGCTCCTCAGCAAGGCTGAAAGGAGGGTGTGGAAAGAGGAGC 1695
QY 1457 TGAGGGTTACCTCTCTGTCCCTGGGACACTATCTCTCATCAACAACCCGCCAGTCTTA 1516
Db 1696 TGAGGGTTACCTCTCTGTCCCTGGGACACTATCTCTCATCAACAACCCGCCAGTCTTA 1755
QY 1517 GAGAGCCACCCCGTACAGAGCCCGGCTCGTGGGAATCCGCCCACTCCGCTCCCT 1576
Db 1756 GAGAGCCACCCCGTACAGAGCCCGGCTCGTGGGAATCCGCCCACTCCGCTCCCT 1815
QY 1577 GTGTCCCAATGGCTCTGCTCTCTCAATCCAGCTTACCCTCTCTTCTGGCCA 1636
Db 1816 GTGTCCCAATGGCTCTGCTCTCTCAATCCAGCTTACCCTCTCTTCTGGCCA 1875
QY 1637 CTTACGCGGCTCCCTTCGAGGCGCGGCCCCCCCCACACCCGCTGGGCCAAACCCACCA 1696
Db 1876 CTTACGCGGCTCCCTTCGAGGCGCGGCCCCCCCCACACCCGCTGGGCCAAACCCACCA 1935
QY 1697 ACACCCAGGCTACAGTGGGACTATATGAGGCTGAGAAAGCCAGGCGCCGCTTCTGC 1756
Db 1936 ACACCCAGGCTACAGTGGGACTATATGAGGCTGAGAAAGCCAGGCGCCGCTTCTGC 1995
QY 1757 CCCCACCTCCCAAGACAGCGTCCCCATATGCCAGGCTGACATTTGTTACCTTCAGG 1816
Db 1996 CCCCACCTCCCAAGACAGCGTCCCCATATGCCAGGCTGACATTTGTTACCTTCAGG 2055
QY 1817 GCGTCACCGGGGCAACACCTATGCTGTGCTGCCTGCACTGCCCCAGGGGAGTCGGGATG 1876
Db 2056 GCGTCACCGGGGCAACACCTATGCTGTGCTGCCTGCACTGCCCCAGGGGAGTCGGGATG 2115
QY 1877 GCGCCCCCAGAGTGGATTTCCCTCGATCTCGACTCCGCTTCAAGAGAGAGCTTGGCGAGG 1936
Db 2116 GCGCCCCCAGAGTGGATTTCCCTCGATCTCGACTCCGCTTCAAGAGAGAGCTTGGCGAGG 2175
QY 1937 GCCAGTTTGGGAGGTGCACCTGTGTGAGTGCAGAGCCCTCAAGATCTGGTCAGTCTTG 1996
Db 2176 GCCAGTTTGGGAGGTGCACCTGTGTGAGTGCAGAGCCCTCAACATCTGGTCAGTCTTG 2235
QY 1997 ATTTCCCTTAAATGCGTAAGGACACCTTTGCTGGTAGCTGTCAAGATCTTACGGC 2056
Db 2236 ATTTCCCTTAAATGCGTAAGGACACCTTTGCTGGTAGCTGTCAAGATCTTACGGC 2295
QY 2057 CAGATGCCACCAAGATG-----CCAGGAATGATTTCTCTGAAAGAGG 2098
Db 2296 CAGATGCCACCAAGATGCCAGCTTCTCTTGTCTCCAGGAATGATTTCTCTGAAAGAGG 2355
QY 2099 TGAAGATCATGTGAGGCTCAAGGACCCCAACATCATTCGGCTGTGGCGGTGTGTGTGTC 2158
Db 2356 TGAAGATCATGTGAGGCTCAAGGACCCCAACATCATTCGGCTGTGGCGGTGTGTGTGTC 2415
QY 2159 AGGACGACCCCTCTGATGATTTACTTACATGAGAGACCGGACCTCAACAGTCTTC 2218
Db 2416 AGGACGACCCCTCTGATGATTTACTTACATGAGAGACCGGACCTCAACAGTCTTC 2475
QY 2219 TCAGTGCCACACAGCTGGAGGACAAAGGACCCGAGGGGGCCCTGGGGAGCGGACGCTG 2278
Db 2476 TCAGTGCCACACAGCTGGAGGACAAAGGACCCGAGGGGGCCCTGGGGAGCGGACGCTG 2535

Db	2549	ATGGCAACCCAGATTCCCTCTGGTATGAAGTACCTTTTCGTCTCAACTTTGTCCACCGA	2508
Qy	2377	GACCTGGCCACCGGAACTSCCTAGTTGGGAAAAATTTACCATCAAAATCGCAGACTTT	2436
Db	2609	GATCTGCCACACAGAACTGTTTGTAGTGGCAAGAATTACACATCAAGATAGCTGATTTT	2668
Qy	2437	GGCATGAGCCGGAACCTCTATCCTGGGACATTAACCTGTCTGGAGGCCGGGCAAGTCTG	2496
Db	2669	GGCATGAGCAGAACTCTGACAGTGGTGAATACTACCGGATCCAGAGGCCGGGGTGTCTC	2728
Qy	2497	CCCATCCGTGGATGSCCTGGGAGTGCATCTCTCATGSGGGAAGTTTCAGAGTGCAGATGCAC	2556
Db	2729	CCCATTCGTGGATGTCCTGGAAAGCATCTTGTGGGCAAAATTCACCGGCAAGTGAT	2788
Qy	2557	GTCTGGCCCTTTGGTGTGACCTCTGGGAGGTGCTGATGCTCTGTAGGGCCAGCCCTTT	2616
Db	2789	GTGTGGCCCTTTGGGTGTACTCTGTGGAGACCTTTCACCTTTTCCAGAGACAGCCCTAT	2848
Qy	2617	GGCAGCTCACCAGCAGCAGAGCTCATCGAGAACGGGGGAGTTCTTCCGGGACACAGGC	2676
Db	2849	TCCAGCTGTCSGATGAGCAGGTTATCGAGAACACTGGAGAGTTCTTCCGAGACCAAGG	2908
Qy	2677	CGCAGGTGTACTGTCCCGGCCCTCGCTTCCCGCAGGSCCTATATGAGCTGATGCTT	2736
Db	2909	AGCAGATCTATCTCCCTCAACACAGCCCTTTGCCCGACTCTGTGTATAAGACTGATGCTC	2968
Qy	2737	CGGTGCTGGAGCCGGGAGTCTGAGCAGCGACACCCCTTTTCCCAAGCTGCATCGTTCCTG	2796
Db	2969	AGCTGCTGGAGAGAGAAACCAAGCACCGGCATCTTCCAGGAAATACACCTCCTGCTT	3028
Qy	2797	GCAGAGATGCATCAACACGGTGTGAATCA	2827
Db	3029	CTTCACAGAGCCGAGTGATGATGATCA	3059

RESULT 6

```

V65317
ID V65317 standard; DNA; 3120 BP.
AC V65317;
DE 22-JAN-1999 (first entry)
DE Receptor protein tyrosine kinase (PTK) subtype tyro-10 encoding DNA.
KW PTK; receptor; protein tyrosine kinase; recombinant; grafting;
KW diagnosis; tumour; skin transplant; connective tissue; tyro-10; ss.
OS Rattus sp.
FT Key Location/Qualifiers
FT 485..3049
FT CDS /*tag= a
/*product= "PTK subtype tyro-10"
US5837448-A.
PD 17-NOV-1998.
AC V65317;
DE 02-MAY-1994; 237401.
DE 15-MAY-1992; US-884486.
DE 02-MAY-1994; US-237401.
DE (SALK ) SALK INST BIOLOGICAL STUDIES.
PA Lai CHC, Lemke GE;
PI WPI; 99-023436/02.
DR P-PSDB; W81409.
PT Nucleic acids encoding protein tyrosine kinase sub:types - for
PT identification of new sub:types and treatment of diseases associated
PT with the kinase
PS Claim 1; Columns 53-58; 47pp; English.
CC This DNA encodes a receptor protein tyrosine kinase (PTK) subtype
CC tyro-10. The invention provides sequences V65308 to V65313, V65315, and
CC V65317 to V65319 that encode proteins having a tyrosine kinase domain and
CC a tissue expression pattern of a receptor PTK subtype selected from
CC tyro-1, tyro-3, tyro-3, tyro-4, tyro-5, tyro-6, tyro-8, tyro-10, tyro-11,
CC and tyro-12, respectively. The polynucleotides are useful for the
CC detection of tyrosine kinase domain sequences and detection of tissue
CC expression patterns of PTK subtypes. The cDNAs can also be injected into
CC oocytes, the protein expressed, and expression products screened for
CC using antibodies against tyrosine kinase epitopes. These subtypes
CC sequences can be used for the design of oligonucleotides, for use in
CC amplification reactions to isolate other subtype sequences. These

```


QY 1000 CGTGGCCCTGCTGCTGGAGGGGAGCCCATCGGCCAACAACTAGGGGGCAACCTG 1059
 Db 1406 TCGG---AAGCCAGGAGTGGGAACCACTGCTGCTACTTTCCCTGCTGCTGACGAT 1462
 QY 1060 GGGAGCCAGAGCCGGGCTGCTCAGTGGCCCTTGGGGGGGCTGCTGCTGCTGCTG 1119
 Db 1463 GTGAACCCAGTCCCGGGTTGTCAGGGTGGCCCTCACCCAGATGGCCAGTGGCCATC 1522
 QY 1120 CAGTGGCTTCTCTTTGGGGGGCTGTTACTCTCTCAGGAAATCTCTTCTATCTCT 1179
 Db 1523 AGTGCCAAATACCATTTTGGCCAGACGCTGGATGATGTTTCCAGGAGATCACTTCCAAATCA 1582
 QY 1180 GATGTGTGAACAAATCTCTCTCGGACCTGGGAGGACCTTCCCGCCAGCCCTGGTGG 1239
 Db 1583 GATGCTCAATGTATAACAACTCTGGAGCCCTTCCCACTCTCTTA----- 1629
 QY 1240 CCGCTGGCCCACTCCCACTTTCAGCAGCTTGGAGCTTGGAGCCAGAGCCAGCAG 1299
 Db 1629 -----TGGCACCCACCACTATGAT 1648
 QY 1300 CCGTGGCCCAAGCCGAGGGAGCCGACCGCCATCTCATCGCTGCTGCTGGCCATC 1359
 Db 1649 CCAATGCTTAAAGTGTATGATACCACTCGGATCTGATTTGGTGGTGGCCATC 1708
 QY 1360 ATCTGCTCTGCTGCTATCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1419
 Db 1709 ATCTTCACTCTGCTGCTATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1768
 QY 1420 CTGAGCAGCTGAGCGGGTGTGGAGAGGAGCTGACGCTTCCACCTCTGCTGCT 1479
 Db 1769 CTAGAAAGGCTTACGAGGAGTGTGGATGATGAATGACATGACCTTCCCTGCTG 1828
 QY 1480 GGGGACACTTCTCTCAACAAACCGCCAGGCTCTAGAGAGCCACCCCGCTACAGGAG 1539
 Db 1829 AGCAGTCCAGCATGTTCAATACACCGCTCTCT-----ATCACCAAGTGAACAGAG 1882
 QY 1540 CCGGGGCTCGTGGGAATCGCCCACTCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1599
 Db 1883 TCAACTCTACTTATGATGAATCTTCCCTTCCCTGCTGCTGCTGCTGCTGCTGCT 1942
 QY 1600 CTGCTCTCAATCCAGCTACCGCTCTCTTCTGCTGCTGCTGCTGCTGCTGCTG 1659
 Db 1943 CTGATC----- 1949
 QY 1660 CCGGGCCCGCCACACCGCTGGGCAAAACCCACCAACCCAGGCTTACAGTGGGAG 1719
 Db 1949 -----CGAAGCTTCCAGAGTTTCTCCAGGAGAGGAGGA 1983
 QY 1720 TATATGAGCGCTGAGAGCGGCGGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 1779
 Db 1984 GTGAGGTGAGTGGTGTGTTGAAGCGGCG-----CAGCCCAATGGACCTGAGGGCGTG 2038
 QY 1780 CCGCATATCGGAGGCTGACATGTTTACCTCGAGGGCTCACCGGGGCAACACCTAT 1839
 Db 2039 CCGCATATCGGAGGCGGAGTGAATCTCCAGGAGTGACAGTGGCAACACCTATC 2098
 QY 1840 GCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1896
 Db 2099 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2158
 QY 1897 CCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1956
 Db 2159 CCGAGGAACTGTTGGGCTTCAAGGAGAGCTGGGAGAGGCGGCTTGGGGAGGCTCAT 2218
 QY 1957 CTGTGTGAGGCTGACAGCCCTCAAGATCTGCTGCTGCTGCTGCTGCTGCTGCTG 2016
 Db 2219 CTGTGTGAGGCTGAGGAGTGAAGAAATTCAGGAGAAAGATTTTGGCACTGATGCTAGT 2278
 QY 2017 AAGGACACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2076
 Db 2279 GCGAACCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2338
 QY 2077 AGGAATGATTTCTGAAAGAGGTGAAGATCATGCTGAGGCTCAAGGACCCCAACATCAT 2136

Db 2339 AGGAATGATTTCTTAAGGAGATCAAGATCATGCTGCTGCTCAAGGACCCAAACATCATC 2398
 QY 2137 CGGCTGCTGGCGGTGTGTGAGGAGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2196
 Db 2399 CGTCTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2458
 QY 2197 AACGGGACCTCAACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2256
 Db 2459 AATGGAGATCTTAATCAGTTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2502
 QY 2257 GCCCTGGGAGGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2316
 Db 2502 -----GTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2548
 QY 2317 GTGCGAGCCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2376
 Db 2549 ATGCAACCCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2608
 QY 2377 GACCTGCGCAGCGGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2436
 Db 2609 GATCTGGCCACACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2668
 QY 2437 GGCATGAGCGGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2496
 Db 2669 GGCATGAGCAGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2728
 QY 2497 CCCATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2556
 Db 2729 CCCATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2788
 QY 2557 GTGTGGGCTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2616
 Db 2789 GTGTGGGCTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2848
 QY 2617 GGGCAGCTCACCGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2676
 Db 2849 TCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2908
 QY 2677 CGGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2736
 Db 2909 AGGAGATCTATCTCTCTCAACAGCCCTTGGCCGCTGCTGCTGCTGCTGCTGCTGCTGCT 2968
 QY 2737 CGGTGCTGGAGCGGAGTCTGAGCAGGAGCCCTTTCCTGCTGCTGCTGCTGCTGCTGCTGCT 2796
 Db 2969 AGCTGCTGGAGAGAGAAACCAAGCAGCGGCTTCTCCAGGAAATACACCTCCTGCTT 3028
 QY 2797 GCAGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2827
 Db 3029 CTTGAGCAAGGAGCGGAGTGTGATGATCA 3059

RESULT 7

Q92521 standard; cDNA to mRNA; 3157 BP.
 AC Q92521;
 DE 26-NOV-1995 (first entry)
 KW Human colonic adenocarcinoma kinase 2 (CK-2) cDNA.
 KW Mammary carcinoma kinase 10; MCK-10; transmembrane receptor; CK-2;
 OS receptor tyrosine kinase; colonic adenocarcinoma kinase 2; cancer; ss.
 FH Homo sapiens.
 FT Key: Location/Qualifiers
 cds 370..2937
 FT /*tag= a
 PN W09514088-A.
 PD 26-MAY-1995.
 PF 16-NOV-1994; E03797.
 PR 16-NOV-1993; US-153397.
 PA (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
 PI Alves FHE, Ulrich A;
 DR WPI; 95-224054/29.
 DR P-PSDB; R75503.
 PT New nucleic acid encoding MCK-10 receptor tyrosine kinase - and

QY 1958 TGTGTAGGTCGACAGCCCTCAAGATCTGGTCACTCTTGATTTCCCTTAAATGTCGGA 2017
Db 2105 TCTGTGAAGTGGAGGAATGGAATAATCAAGACAAGATTTTCCCTAGATCTCAGTG 2164
QY 2018 AGGACACCCCTTCTGGTAGCTGCAAGATCTTACGGCCAGATGCCACCAAGAATGCCA 2077
Db 2165 CAACACAGCTGCTCTGGTGGCTGTGAATAATGCTCCGAGCAGATGCCACAAGAATGCCA 2224
QY 2078 GGAATGATTTTCTTGAAGAGGTGAAGATCATGTCTGAGGCTCAAGGCCCAACATCATTC 2137
Db 2225 GGAATGATTTTCTTGAAGAGATAAGATCATGTCTGAGCTCAAGGCCCAACATCATTC 2284
QY 2138 GGCTGTGGGCGTGTGTGACGACGACCCCTCTGCTCATGATTAATGATCATGAGAGA 2197
Db 2285 ATCTATTATCTGTGTGATCACTGATGACCTCTCTGTATGATCACTGAATATACATGAGA 2344
QY 2198 ACGGCGACCTCAACAGTCTCTCAGTCCGACAGCTGGAGGACAAGGCGAGGGGG 2257
Db 2345 ATGAGATCTCAATCAGTTCTTCCGCGCACGAGCCCTTAATCTCTCCAGCG--- 2402
QY 2258 CCCCTGGGCGGCGAGGCTCGGACGGGGCCACCATCATCAGCTACCCAAATGCTGTCATG 2317
Db 2402 -----ATGACGCACTGTCACTTACACCAATCTGAATTTA 2437
QY 2318 TGGCAGCCAGATCGCTCGGCGATGCGCTATCTGGCCACATCAACTTTGTATACGCGG 2377
Db 2438 TGGCTACCAAAATGCTCTGGCATGAATGACCTTTCTCTCTTAATTTTGTTCACCGAG 2497
QY 2378 ACCTGGCCACCGGACTGCTAGTCTGGGAAATTTTACCACCAAAATCGACAGCTTTG 2437
Db 2498 ATCTGGCCACACGAACTGTTTATGTTGGTAAGAACTACAACTAAGATGACTGTTG 2557
QY 2438 GCATGAGCGGAACTCTATCTGGGAGTATTACCTGTGCGAGGGCGGCGAGTGTGTC 2497
Db 2558 GAATGAGCAGCACTGTACAGTGTGACATTTACCAGATCCAGGGCGGCGAGTGTGTC 2617
QY 2498 CCATCCGCTGATGCTGGGAGTGCATCTCTATGCGGAGTTCACGACTGCGAGTGAAG 2557
Db 2618 CTATCCGCTGATGCTTGGGAGATGCTTCTGCTGGCAAGTTTCACTACACGAACTGATG 2677
QY 2558 TGTGGCCTTTGGTGTGACCTGTGGAGGTGCTGATGCTCTGTAGGCCCGACCCCTTTG 2617
Db 2678 TGTGGCCTTTGGGTTACTTTGTGGGAGTCTTACCTTTTGTCAAGAACAGCCCTATT 2737
QY 2618 GGCAGCTCACCAGCAGCAGTCTATCGAAGACGCGGGGAGTCTTCCGGGACAGGCGC 2677
Db 2738 CCCAGCTGTGATGACAGCTTATTGAGATACTGAGAGTCTTCCGAGACCAAGGA 2797
QY 2678 GGCAGGTGATCTGTCCGGCGGCTGCTGCTCCGCGAGGGCTTATATGAGCTGATGCTTC 2737
Db 2798 GGCAGACTTACCTCCCTCAACAGCCATTTGTCTGACTCTGTGTATGATGATGCTCA 2857
QY 2738 GGTGCTGGAGCGGGAGTCTGAGCAGCGACCAACCTTTTCCAGCTGCATCGGTTCT 2795
Db 2858 GCTGCTGGAGAGAGATACGAAAGACCGTCCCTCATTTCCAAAGAAATCCACTTCTGCT 2915

RESULT 8

ID Q92523 standard; cDNA to mRNA; 3157 BP.
AC Q92523;
DT 26-NOV-1995 (first entry)
DE Human colonic adenocarcinoma kinase 2 (CK-2) cDNA.
KW Mammary carcinoma kinase 10; MCK-10; transmembrane receptor; CK-2;
KW receptor tyrosine kinase; colonic adenocarcinoma kinase 2; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT cds 370..2937
FT /*tag= a
PN WO9514089-A.
PD 26-MAY-1995.
PF 16-NOV-1994; E03799.
PR 16-NOV-1993; US-153397.

PA (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
PI Alves.FHE, Ullrich A;
DR WPI; 95-224055/29.
DR P-PSDB; R75505.
PT New nucleic acid encoding CK-2 receptor tyrosine kinase - and
PT derived vectors, transformed cells, proteins and antibodies, useful
PT for diagnosis and treatment of proliferative and nervous system
PT diseases and for screening modulators
PS Disclosure; Page 73-74; 115pp; English.
PS A member of the mammary carcinoma kinase 10 (MCK-10) receptor
CC tyrosine kinase family was identified using a PCR (with two
CC degenerate oligo primer pools based on conserved sequences of the
CC kinase domains of receptor tyrosine kinases) and cDNA prepd. from
CC colonic adenocarcinoma RNA. The nt sequence of the novel receptor,
CC designated CK-2, is given in Q92523 and the deduced AA sequence in
CC R75505. Analysis of CK-2 nt and AA sequences indicated significant
CC homology with MCK-10 throughout the extracellular, transmembrane
CC and intracellular regions. The regions of homology extend into the
CC N-terminus consensus sequence of the discoidin I like family of
CC proteins. CK-2 was predominantly found in all stromal cells
CC whereas MCK-10 expression was strongly confined to neoplastic
CC cells themselves. Between the two RTKs, the juxtamembrane region
CC is the region of most extensive sequence divergence. 806 T;
SQ Sequence 3157 BP; 790 A; 802 C; 759 G;

Query Match 18.3%; Score 666.8; DB 1; Length 3157;
Best Local Similarity 56.8%; Pred. No. 5e-127;
Matches 1532; Conservative 0; Mismatches 977; Indels 189; Gaps 9;

QY 110 TGCTGCTGCTTTGGTGGCAAGTGGAGATGCTGACATGAAGGGACATTTTGTATCCTGCCA 169
Db 395 TGGTCTGCTTCTGCTGCTGCTATCTTGTGCTTCTGCAAAAGCTCAGGTTAATCAGCTA 454
QY 170 AGTGGCGGTATGCTGCTGGCATGACAGCCGACCATCCAGACAGTACATCTCTGCTT 229
Db 455 ATGTCGGGTATCTCTCTGGCATGTCAGGAGCCAGATTCCAGATGAGGACATCAGCTT 514
QY 230 CCAGCTCTCTGCTGACATCTCCACTGCCGCCGCCAGCAGGTTGGAGAGCAGTACGCGG 289
Db 515 CCAGTCAGTGGTTCAGATCCAGCTCCAGCTGCCAAATATGGAAGCTGGACTCAGAAGAAG 574
QY 290 ATGGGCGCTGTGTCGCCGAGGCTCGGTGTTTCCCAAGGA---GGAGGAGTACTTGCAG 346
Db 575 ATGGAGCCTGTGTCCTGAGATTCAGTGGAACTGATGACCTGAAGAGGTTTCTGCAGA 634
QY 347 TGSATCTCAACAGCTGCACCTGGTGGCTGCTGGTGGGACCCAGGAGCGGATCCGCGG 406
Db 635 TTGACTTGCACACCTCCATTTTATCACTCTGGTGGGACCCAGGCGGCCCATGCGAGAG 694
QY 407 GCCTGGCAAGAGTCTCTCCGGAGCTACCGGCTGCTTACTCCCGGATGCTGCGCGCT 466
Db 695 GTCATGGCATCAGTTTGCCTCCATGTACAAGATCAATTACAGTCGGGATGSCACTCGCT 754
QY 467 GGATGGCTGGAAGACCGCTGGGGTCAAGAGGTGATCTCAGGCAATGAGGACCTGAGG 526
Db 755 GGATCTCTTGGCGGAACCGTCTATGGGAAACAGGTGCTGGATGGAATAGTAACCCCTATG 814
QY 527 GAGTGTGCTGAAGACCTTGGGCCCCCATGTTGGCCCGACTGGTTCCTTACCTTACCCCC 586
Db 815 ACATTTTCTTAAGGACTTGGAGCGCCCATTTGTAGCCAGATTTTCCGGTTTATTCCAG 874
QY 587 GGGCTGACCGGCTCATGACGCTCTCTCTCGCGGTAGAGCTCTATGGCTCCCTCGGAGG 646
Db 875 TCACCGACCACTCCATGAATGTGTATGATGAGAGTGGAGCTTTACGGCTGTGCTGGCTAG 934
QY 647 ATGGACTCTCTTACACCGCCCTGTGGGCGACACAATGTATTTATCTGAGG----- 701
Db 935 ATGGCTTGGTGTCTTACAATGCTCCAGCTGGGCGAGCTTTGTACTCCTCGAGGTTCCA 994
QY 701 CCGTGTACCTCAACGACTCCACCTATGACGACATACCTGGGCGGACTGCGAGTATGGG 760
Db 995 TCATTTATCTGAATGATCTGTCTATGATGAGCTGTTGGATACAGCATGACAGAGGG- 1054

QY 761 GTCTGGCCAGCTGGCAGATGTTGGTGGGCTGGATGACTTTAGGAAGATCAGGAGC 820.
Db 1054 --CTAGGCCAATTGACCGATGTTGCTGGCTGGAGATTTCACCGAGCCCATGAAT 1111
QY 821 TGGCGGTCTGGCCAGGCTATGACTATGTTGGGATGAGAACACACAGCTTCTCCAGTGGCT 880
Db 1112 ACCACGTGTGGCCGGCTATGACTATGTTGGGCTGGGAAACAGAGTGCACCAATGGCT 1171
QY 881 ATGTGAGATGGAGTTGAGTTGACCGGCTGAGGGCTTCCAGGCTATGAGGTTCCACT 940
Db 1172 ACATTGAGATCATGTTGAAATTTGACCGCATCAGGAATTTCACTACCATGAAGTTCCTACT 1231
QY 941 GTAACAACATGACACAGCTGGGAGGCCGCTCTCCCTGGGGGTGAAATGTCGCTTCCGGC 1000
Db 1232 GCAACAACATGTTGCTAAAGGTGAAGATCTTTAAGGAGGTACAGTGTACTTCCGCT 1291
QY 1001 GTGGCCCTGCCATGCTGGGAGGGGAGCCCATGGCCACAACTAGGGGGCAACCTGG 1060
Db 1292 CTG---AAGCCAGTGAAGTGGGAACCTAATGCCATTTCCCTTCCCTTCTGCGATGAGC 1348
QY 1061 GGGACCCAGAGCCGGGCTGTCTCAGTGGCCCTTGGGGCGGTGGCTGCTGCTTTCTGC 1120
Db 1349 TCAACCCAGTGTGGTGTGTCAGGGTGCCTCTCCACCAGGATGGCGAGTGCCTATCA 1408
QY 1121 AGTGGCGTCTCTTGTGGGGGCTCTGGTACTCTCTCAGCGAAATCTCTTCACTCTCTG 1180
Db 1409 AGTGTCAATACCATTTGTCAGATACCTGGATGATGTTCACTGAGATCACTTCCCAATCAG 1468
QY 1181 ATCTGTGAACAATTCCTCTCGGCACATGGGAGGACCTTCCGCCAGCCGCTGGTGGC 1240
Db 1469 ATGCTCAATGTACAACAACCTGTAAGCCCTGCCCACTCTCTCTA----- 1514
QY 1241 CGCCTGGCCCACTCTCCCACTTTCAGCAGCTTGGAGCTGGAGCCAGAGCCAGCAGC 1300
Db 1514 -----TGACCCCAACCTATGATC 1534
QY 1301 CCGTGGCCAAAGCCGAGGGAGCCGAGCCGATCCTCATCGGCTGCTGTGGCCATCA 1360
Db 1535 CAATGCTTAAAGTTGATGACAGCAACACTCGGATCCTGATTGGCTGTGTTGGTGGCATCA 1594
QY 1361 TCTGCTCTCTGCTCATCATGTTGCCCTCATGCTCTGGGGGTGACATGGGCGAGGCTCC 1420
Db 1595 TCTTATCTCTCTGCGCATCATGTTGATCATCTCTGAGCGAGTCTGGCAGAAATGTC 1654
QY 1421 TCAGCAAGGCTGAAGGAGGTGTTGGAAGAGAGCTGACGCTTCACTCTCTGTCCTG 1480
Db 1655 TGGAGAAGGCTTCTCGGAGATGCTGGATGATGAATGACAGTCAAGCTTTCCTGCGCAA 1714
QY 1481 GGGACATATCTCATCAACACCCGCGAGTCTCTAGAGGCCACCCCGTACCAGGAGC 1540
Db 1715 GTGATTAGCATGTTCAACAAT-----AACCGCTCCT 1747
QY 1541 CCGGCTCTGGGATCGCCCACTCGCTCCCTGCTGCTCCCAATGGCTCTGGGTTGC 1600
Db 1748 CATCACTAGTGAACAAGGTCACATGCTGATGATGCTGCTGCTGCTGCTGCTGCTGCT 1795
QY 1601 TGCCTTCCAATCCAGCTACCCGCTCTCTGCGCCACTTACGCGCTCCGCTCGAGGCC 1660
Db 1795 -----CCCTTCCGCT 1804
QY 1661 CGGGCCCCCAACCCGCTGGGCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1720
Db 1805 CTGACTTACCAGGAGGCTACCGGCTGATACGAAACTCCAGAAATTTGCTCCAGGGAGG 1864
QY 1721 ATATGAGGCTGAGAGCCAGGCGCCGCTCTGCGCCGCTCCGCTCCGCTCCGCTCCGCT 1780
Db 1865 AGAGTCAAGGCTGAGCGGTGTTGAAGCCAGTCTCCAGCCAGTGGGCTGAGGGGTGC 1924
QY 1781 CCATTATCCGAGGCTGACATTTGTAACCTGACGGGTGTCACCGGGGCAACACTATG 1840
Db 1925 CCCACTATCGAGGCTGACATGTAACCTGACGGGTGTCACCGGGGCAACACTATG 1984

QY 1841 CTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1897
Db 1985 CAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2044
QY 1988 CTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1957
Db 2045 CCAGGAACTCTTAACTTTTCAAGAGAAAGTGGGAGAAAGTGGGAGGTTTGGGAGGTTTCATC 2104
QY 1958 TGTGTGAGTGCAGACGCCCTCAAGATCTGGTCACTTGTGATTTCCCTTAAATGTCGTA 2017
Db 2105 TCTGTGAAGTGGAGGAATGGAATAATCAAGACAAAGATTTTCCCTAGATGTCAAGT 2164
QY 2018 AGGACACCCCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2077
Db 2165 CCAACAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2224
QY 2078 GGAATGATTTCTGAAAGAGTGAAGATCATGTCGAGGCTCAAGGACCCCAACATCATTC 2137
Db 2225 GGAATGATTTCTTAAAGAGATAAAGATCATGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2284
QY 2138 GGTGCTGGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2197
Db 2285 ATCTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2344
QY 2198 ACGGCAGCTCAACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2257
Db 2345 ATGGAGCTCAATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2402
QY 2258 CCGCTGGGAGGGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2317
Db 2402 -----ATGTAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2437
QY 2318 TGGCAGCCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2377
Db 2438 TGGTACCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2497
QY 2378 ACCTGGCAGCGGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2437
Db 2498 ATCTGGCCACACCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2557
QY 2438 GCATGCGCGGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2497
Db 2558 GAATGAGCAGGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2617
QY 2498 CCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2557
Db 2618 CTATCCCTGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2677
QY 2558 TGTGGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2617
Db 2678 TGTGGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2737
QY 2618 GGCAGCTCACCAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2677
Db 2738 CCAGCT 2797
QY 2678 GGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2737
Db 2798 GGCAGATTTACCTCCCTCAACAGCCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2857
QY 2738 GGTGCTGGAGCGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2795
Db 2858 GCTGCTGGAGAGAGATACGAGAACCGTCCCTCATTCACGAANAATCCACCTTCTGCT 2915

RESULT 9

T93784

ID T93784 standard; cDNA; 3157 bp.

AC T93784;

DT 16-FEB-1998 (first entry)

DE CCK-2, a human mammary carcinoma kinase 10 MCK-10) family member, cDNA.

KW Mammary carcinoma kinase; MCK-10; receptor tyrosine kinase;

proliferative disease; cancer; MCK-10 activity; aberrant expression; ds.

OS Homo sapiens. Location/Qualifiers
 FH Key 370..2934
 FT CDS /*tag= a
 PN US5677144-A.
 PD 14-OCT-1997. 336343.
 PF 08-NOV-1994; 336343.
 PR 16-NOV-1993; US-153397.
 PA (ALVE/) ALVES F H E.
 PA (ULLR/) ULLRICH A.
 PI Alves FHE, Ullrich A;
 DR WFL: 97-511869/47.
 DR P-PSDB: W34671.
 PT Truncated receptor tyrosine kinase CKK-2 - and nucleic acid coding
 PR for it, useful for cancer diagnosis
 PS Claim 9; Fig 3; 70pp; English.
 CC The present sequence represents the cDNA sequence of human CKK-2, a
 CC member of the mammalian tyrosine kinase 10 (MCK-10, T93785) family of
 CC receptor tyrosine kinases. Expression of CKK-2 is associated with
 CC proliferative diseases such as cancer. The CKK-2 gene was identified by
 CC PCR and a cDNA prepared from colon adenocarcinoma RNA. CKK-2 is
 CC expressed in a wide variety of cancer cell lines and tumour tissue. The
 CC CKK-2 nucleic acids can be used for diagnostic purposes to detect
 CC aberrant expression of CKK-2 genes. Engineered cell lines, containing
 CC recombinant vectors with the present sequence, are useful for producing
 CC infectious retroviral particles. The cell lines may also be used to
 CC evaluate and screen drugs involved in CKK-2 activation and regulation.
 SQ Sequence 3157 BP; 790 A; 802 C; 759 G; 806 T;

Query Match 18.3%; Score 666.8; DB 1; Length 3157;
 Best Local Similarity 56.8%; Pred. No. 5e-127;
 Matches 1532; Conservative 0; Mismatches 977; Indels 189; Gaps 9;

QY 110 TGCTGCTGCTCTGGTGGCAAGTGGAGATGCTGACATGAAGGACATTTTGATCCGCA 169
 DB 395 TGGTGTGTTCTCTGCTGCTGCTATCTTTAGTCTGCAAAAGCTCAGGTTAATCCAGTA 454
 QY 170 AGTCCGCTATGCTGGGATGAGGACGAGGACATCCAGACAGTGCATCTCTGCTT 229
 DB 455 TATCCGCTATCTCTGGGATGTCAGGAGGCGGAGATTCAGATGAGGACATCACAGCT 514
 QY 230 CCAGCTCTGTGATGATTCAGTCCGCGCCGCGGACGAGTGGAGAGAGTGCAGGGG 289
 DB 515 CCAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 574
 QY 290 ATGGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 346
 DB 575 ATGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 634
 QY 347 TGGATCTACACGACTGCACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 406
 DB 635 TTGACTTGACACCTCCATTTTATCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 694
 QY 407 CCTGTTGGAAGAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 466
 DB 695 GTATGGCATCGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 754
 QY 467 GGATGGCTGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 526
 DB 755 GGATCTCTTGGCGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 814
 QY 527 GAGTGGTCTGAAGAGCTTTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 586
 DB 815 ACATTTTCTTAAGGAGCTTGGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 874
 QY 587 GGGTGTACCGGGTCTATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 646
 DB 875 TCACCGACCACTCCATGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 934
 QY 647 ATGACCTCTCTTACACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 701
 DB 935 ATGCTTGGTCTTACAACTCTCCAGCTGGGAGAGAGTGTGATCTCCCTTGGAGGTTC 994

QY 701 CCGTGTACCTCAACGACTCCACCTATGACGGACATACCGTGGCGGACTGCAGTATGGG 760
 DB 995 TCATTTATCTGAATGATTCTGTCTATGATGAGCTGTTGGATACAGCATGACAGAGGG- 1054
 QY 761 GTCGTGGCCAGCTGGCAGATGGTGTGGTGGGCTGGATGACTTTAGGAAGAGTGCAGGAGC 820
 DB 1054 --CTAGGCCAATTGACCGATGGTGTGTGCTGGCCTGGAGATTTTCCACGACACCATGAAT 1111
 QY 821 TCGGGGCTCGCCAGGCTATGACTATGTTGGATGAGCAACACAGCTTCTCCAGTGGCT 880
 DB 1112 ACCAGGTGTGGCGCTGATGACTATGTTGGCTGGCGAAGAGAGTGCACCAATGGCT 1171
 QY 881 ATGTGGAGATGAGTGTGAGTTTGACCGGCTGAGGGCTTCTCCAGGCTATGACAGTCCACT 940
 DB 1172 ACATTTGAGATCATGTTTGAATTTGACCGCATCAGAAATTTTCACTTACCATGAAGTCCACT 1231
 QY 941 GTAACAATCATCACAGCTGGGAGCGCTGCTGCTGGGGGTGGAATGTGCTTCCGGC 1000
 DB 1232 GCAACAACATCTTTGCTAAAGGTGTGAAGATCTTTAAAGAGGTACAGTGTCTACTTCCGCT 1291
 QY 1001 GTGGCCCTGCCATGCCCTGGGAGGAGCCCATGCGCCACCACTAGGGGGCAACCTGG 1060
 DB 1292 CTG---AAGCCAGTGTGGAACTATGCAATTTCTTCCCTTGTCTCTGGATGAGC 1348
 QY 1061 GGGACCCAGAGCCCGGCTGCTCTAGTGGCCCTTGGCGGCGTGTGGCTGCTTCTGCTG 1120
 DB 1349 TCAACCCAGTGTCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1408
 QY 1121 AGTGGCGTCTCTTTCGGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1180
 DB 1409 AGTGTCAATATACATTTTGCAGATACCTGGATGATGTTTCACTGAGATCACCTTCCAATCAG 1468
 QY 1181 ATGTGGTCAACAATTCCTCTCCGCGCTGAGGAGCACCTTCCCGCCAGCCCTTGGTGGC 1240
 DB 1469 ATGTGCAATGTACAACTCTGAGCCCTGCGCCACTCTCTTA----- 1514
 QY 1241 GCGCTGGCCCACTCCCAACCTTACAGACTTGGAGCTGGAGCCAGAGCCAGCAGC 1300
 DB 1514 -----TGGCCACCAACCTTATGATC 1534
 QY 1301 CCGTGGCCAAAGCCGAGGGAGCCGACCGCATCTCATCGGCTGCTGCTGCTGCTGCTGCT 1360
 DB 1535 CAATGCTTAAAGTTGATGACAGCAACACTCGGATCTGCTGCTGCTGCTGCTGCTGCTGCT 1594
 QY 1361 TCCTGCTCTGCTGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1420
 DB 1595 TCTTTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1584
 QY 1421 TCAGCAAGCTGAAGGAGGCTTGGAGAGAGCTGACGGTTTCACTCTCTGCTGCTGCTGCT 1480
 DB 1655 TGGAGAAGGCTTCTCGGAGGATGCTGGATGATGAAATGACAGTCACTGCTGCTGCTGCTGCT 1714
 QY 1481 GGGACACTATCTCATCAACAACCGCCAGCTCTAGAGAGCCACCCCTGACAGGAGC 1540
 DB 1715 GTGATTTAGCATGTTCAACAAT-----AACCGCTCT 1747
 QY 1541 CCGGCGCTCGTGGGAATCCGCGCCCACTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1600
 DB 1748 CATCACTAGTGAACAGGGTCCAACTGCTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1795
 QY 1601 TGCTCTCAATCCAGCTTACCGCTCTCTTCTGGGCACTTACGCGGCTGCTGCTGCTGCTGCT 1660
 DB 1795 -----CCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1804
 QY 1661 CGGGCCCCCAGACCCCGCTGGGCGCAACCCACCAACCCAGCCCTACAGTGGGACT 1720
 DB 1805 CTGACTACAGAGGACCTCCAGGCTGATACGAAACTCCCAAGATTTGCTCCAGAGGAGG 1864
 QY 1721 ATATGAGGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1780
 DB 1865 AGGAGTCAAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1924

Db 859 TCACCGACCACTCCATGAATGTGTATGAGAGTGGAGCTTTACGGCTGTCTGGCTAG 918
QY 647 ATGGACTCTCTTACACCGCCCTGTGGGACACAAATGATTTATCTGAGG----- 701
Db 919 ATGGCTTGTGTCTTACATGCTCCAGCTGGGACAGCTTTGTACTCCCTGGAGTTCCA 978
QY 701 CCGTGTACCTAACGACTCCACTATGACGGACATACCGTGGGGAGTGCAGTATGGG 760
Db 979 TCATTTATCTGAATGATCTGTATGATGGAGCTGTGGATACAGCATACAGAAAGG- 1038
QY 761 GTCTGGGACAGCTGCAGATGTGTGTGGGCTGGATGACTTTAGGAAGAGTCAGAGC 820
Db 1038 --CTAGGCCAATGTAGCCGATGTGTCTGGCTGGAGATTTACCCAGACCCATGAAT 1095
QY 821 TGGGGCTGTGGCAGCTATGACTATGTGGATGGAGCAACACAGCTTCTCCAGTGGCT 880
Db 1096 ACCAGCTGTGGCCGCTATGACTATGTGGCTGGCGAAGAGAGTCCACCAATGGCT 1155
QY 881 ATGTGGAGATGAGTTGAGTTTGACCGGCTGAGGGCCTTCCAGGCTATGAGGTCCACT 940
Db 1156 ACATTGAGATCATGTTGAATTTGACCGCATCAGGAATTTCACTACCATGAAGTCCACT 1215
QY 941 GTAAACACATGCACAGCTGGAGCCCTCTGCTGCGGGGTGGAATGTCGCTCCGGC 1000
Db 1216 GCAACACATGTTGCTAAGGTGTGAAGATCTTTAAGGAGGTACAGTGTCTTCCGCT 1275
QY 1001 GTGGCCCTGCCATGGCTGGAGGGAGCCCATCGGCACAACTAGGGGCAACCTGG 1060
Db 1276 CTG---AAGCCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1332
QY 1061 GGAACCCAGAGCCGGGCTGTCTCAAGTGGCCCTTGGGGGGGCTGTGGCTTTCG 1120
Db 1333 TCAACCCAGTGTGGTGTGTGACGGTGTGACGGTGTGACGGTGTGACGGTGTGAC 1392
QY 1121 AGTGGCGCTTCTCTTGGGGGCTGTGTCTTCTGAGGATCTCTTCACTCTG 1180
Db 1393 AGTGTCAATACCAATTTGAGATACCTGTGATGATGTTGAGTGAAGTACCTTCCAA 1452
QY 1181 ATGTGTGAACAATCTCTCGGCACTGGAGGCACTTCCCGCCAGCCCTTGGTGGC 1240
Db 1453 ATGCTCAATGTACACAACTCTGAAGCCCTGCCACCTCTCTA----- 1498
QY 1241 CGCCTGGCCACTCTCCCACTCTGAGAGCTTGGAGCTGGAGCCAGAGCCAGCAGC 1300
Db 1498 -----TGCAACCCACAACTATGATC 1518
QY 1301 CCGTGGCCAGCCGAGGGAGCCGACCGCATCTCATCGCTGCGCTGTGGCCATCA 1360
Db 1519 CAATGCTTAAAGTTGATGACACACACTCGGATCCTGATGGCTGTGTGGCCATCA 1578
QY 1361 TCCTGCTCTGCTCATCTATGCGCTCATCTCTGGCGGCTGCAGCTGGGCGAGCTCC 1420
Db 1579 TCTTTATCTCTGGCCATCATGTCATCTCTCTGGAGGAGTCTGGCAAAATGC 1638
QY 1421 TCAGAAAGGCTGAACGGAGGTGTGGAGAGAGCTGAGGTTTCACTCTGTCCCTG 1480
Db 1639 TGGAGAAAGCTTCTCGAGAGTGTGGATGATGAATGACATGACAGCTTCTCCCTG 1698
QY 1481 GGGACACTCTCTCAACACCCGAGTCTCTGAGAGCCACCCCGGTACAGGAGC 1540
Db 1699 GTATTCTAGCATGTTCAACAAT-----AACCGCTCT 1731
QY 1541 CCGGGCTGTGGGAATCGCCCACTCGCTCCCTGTGTCCCAATGGCTGTGGTTGC 1600
Db 1732 CATCACTAGTGAACAGGTCCAACCTGACCTTACGATCGCATCTT----- 1779
QY 1601 TGTCTCCAATCAGCCTACCGCTCTCTTGGCCACTTACCGCGTCCCGCTCAGGCC 1660
Db 1779 -----CCCTTCGCC 1788
QY 1661 CGGGCCCCCAGACCCGCTGGGCAAAACCCAGCAACCCAGGCTTACAGTGGGACT 1720

Db 1789 CTGACTACAGGAGCCATCCAGGCTGATACGAAACTCCAGAAATTTGCTCCAGGGAGG 1848
QY 1721 ATATGAGGCTTGAGAGCCAGGCGCCCTTCTTGGCCCACTCCCCAGAACAGGTCC 1780
Db 1849 AGGAGTCAAGGCTCAGCGGTGTGTGAAGCCAGTCCAGCCAGTCCCTGAGGGGTGC 1908
QY 1781 CCAATATGCCAGGCTGACATTTTACCTCGAGGCGTCCAGGGGCGACACCTATG 1840
Db 1909 CCACATATGCAGAGGCTGATAGTGAACCTCCAAGAGTGCAGAGGCGACACATCT 1968
QY 1841 CTGTGCTGCACGTGCCCGAGGGGAGTC---GGGATGGGCCCCCAGAGTGGATTC 1897
Db 1969 CAGTGGCTGCCGTACCATGGACCTCTCTCAGAAAGATGTGGCTGTGGAGGATTC 2028
QY 1898 CTCGATCTCGACTCCGCTTCAAGGAGAGCTTGGGAGGCGCAGTTTGGGAGGTGCAC 1957
Db 2029 CCAGGAACTCCTTAACCTTCAAGAGAACTGGGAGAGACAGTTTGGGAGGTTCATC 2088
QY 1958 TGTGTGAGTGCAGACCCCTCAAGATCTGTGAGTCTGTGAGTCTTATTTCCCTTATGTGCGTA 2017
Db 2089 TCTGTGAAGTGGAGGAATGGAAATTTCAAGACAAAGATTTGCCCCAGATGTGAGTG 2148
QY 2018 AGGACACCCCTTGTCTGTAGCTGTCAAGATCTTACGGCAGATCCACCAAGAAATGCCA 2077
Db 2149 CCAACAGCTGTCTGTGGTGTGAAATGCTCGAGCAGATGCCAACAGAAATGCCA 2208
QY 2078 GGAATGATTTCTTAAAGAGGTGAAGATCATGTGAGGCTCAAGGACCCAAACATCATTC 2137
Db 2209 GGAATGATTTCTTAAAGAGATAAGATCATGTCTGGCTCAAGGACCCAAACATCATCC 2268
QY 2138 GGCTGTGGGCTGTGTGAGGAGCGCCCTCTGATGATTTACTGACTACATGGAGA 2197
Db 2269 ATCTATTATCTGTGTATCACTGATGACCCCTCTCTGTATGATCACTGAATACATGGAGA 2328
QY 2198 ACGGCACTTCAACCACTTCTCAGTGCACCCAGCTGGAGGACAAAGGCGAGGGGG 2257
Db 2329 ATGGAGATCTCAATCATGTTCTTCCCGCCAGGAGCCCTTAATTTCTCTCCAGCG--- 2386
QY 2258 CCGCTGGGAGCGGCGAGGCTGCGAGGCGCCACCATCACTAGCTACCCAATGCTGCTCATG 2317
Db 2386 -----ATGTAGCACTGTCACTGATCACTCACTCACTCACTCACTCACTCACT 2421
QY 2318 TGGCAGCCAGATCGCTCCGGCATCGGCTATCTGCGCACACTCACTTTGTACATCGG 2377
Db 2422 TGGCTACCCAATTTGGCTGTGGCATGAACTGCTTCTCTCTTAATTTGTTCACCGAG 2481
QY 2378 ACCTGGCAGCGGAACTGCTGTGTGGGAAATTTTCACTCAAAATTCGCGAGCTTG 2437
Db 2482 ATCTGGCACACGAACTGTTTGTGGTGAAGAACTACAACTCAAGATAGCTGACTTG 2541
QY 2438 GCATGAGCGGAACCTCTATGTGTGGGACTATTTACGCTGTGAGGCGCGGAGTCTGC 2497
Db 2542 GAATGAGCAGGAACCTGTACAGTGTGACTATTTACGGATCCAGGCGCGGAGCTCTCC 2601
QY 2498 CCATCCGCTGGATGGCTGGAGTGCATCTCATGGGAGTTCAGACTGCGAGTGAGC 2557
Db 2602 CTATCCGCTGGATGTCTTGGGAGATCTTGTGTGGGAGTTTCACTACAGAAAGTATG 2661
QY 2558 TGTGGGCTTGTGTGACCTGTGGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2617
Db 2662 TGTGGGCTTGTGGGTACTTTGTGGAGACTTTTACCTTTTGTCAAGAACAGCCCTATT 2721
QY 2618 GGCAGCTCACGAGAGAGGTGATCGAGAACGGGGAGTCTTCCGGGACAGGGCC 2677
Db 2722 CCCAGCTGTGAGATGAACAGGTATTGAGAATACTGGAGAGTTCTTCCGAGCAAGAGGA 2781
QY 2678 GGCAGGTGTACCTGTCCCGCCCTTGCCTGCGGAGGCGCTTATAGCTGATGCTTC 2737
Db 2782 GGCAGACTTACCTCCCTCAACAGCCATTTGTCTGACTCTGTGTATAAGCTGATGCTCA 2841
QY 2738 GGTGCTGGAGCGGAGTCTGAGCAGCGACCCCTTTTCCAGCTGCTCGTTCCT 2795
Db 2842 GCTGCTGGAGAGAGATACCAAGACCGTCCCTCAITCCCAAGAAATCCACCTTCTGCT 2899

RESULT 11

Q64158
ID Q64158 standard; DNA: 2128 BP.
AC Q64158;
DT Q3-FEB-1995 (first entry)
DE Partial coding sequence of tyrosine kinase receptor protein.
KW Tyrosine kinase; receptor; proto-oncogene; trk; detection;
KW diagnosis; antibody; treatment; tumour; antisense; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT 1..1953
FT cds
FT /tag= a
FT /product= Partial sequence of protein tyrosine
FT kinase.
FN DE4239817-A.
FN 01-JUN-1994.
PF 26-NOV-1992; 239817.
PR 26-NOV-1992; DE-239817.
PA (CHEM-) CHEMOTHERAPEUTISCHES FORSCHUNG.
PI Holtrich U, Ruebsamen-waigmann H, Strebhardt K;
DR WPI; 94-184380/23.
DR P-PSDB; R54089.
PT New protein tyrosine kinase and related nucleic acid - vectors,
PT transformed cells, etc., useful for diagnosis and treatment of
PT tumours
PS Claim 4; Page 8; 9pp; German.
CC The gene is related to the trk proto-oncogene. Antibodies against
CC the encoded polypeptide are useful for diagnosis and for the
CC treatment of tumours. The antibodies may also be radiolabelled or
CC coupled to a cytotoxin for destruction of cancer cells. Antisense
CC nucleic acid can be used to inhibit gene expression.
SQ Sequence 2128 BP; 532 A; 556 C; 504 G; 536 T;

Query Match 13.2%; Score 481.4; DB 1; Length 2128;
Best Local Similarity 56.1%; Pred. No. 2.5e-89;
Matches 1176; Conservative 0; Mismatches 741; Indels 180; Gaps 7;

QY 702 CGTGTACCTCAACGACTCCACTATGACGGACATACCGTGGCGGACTGCGAGTATGGGG 761
Db 12 CATTATCTGAATGATCTGTCTAATGATGGAGCTGTGGATACAGCATGACAGAGGG-- 70
QY 762 TCTGGCCAGCTGGCAGATGTTGGTGGGCTGGATGATGATTTAGGAGAGTACAGGAGCT 821
Db 70 -CTAGGCCAATTGACCGATGGTGTCTGTGGCTGGACGATTTCCACCAGACCCATGAATA 128
QY 822 GCGGCTCTGGCCAGCTATGACTATGTGGGATGGAGCAACACAGCTTCTCCAGTGGCTA 881
Db 129 CCACGTGTGGCCGCTATGACTATGTGGGCTGGCGAAGAGAGTGCCCAATGGCTA 188
QY 882 TGTGAGATGGAGTTGAGTTGACCGCTGAGGCGCTTCCAGGCTATGACAGTCCACTG 941
Db 189 CATTGAGATCATGTTGATTTGACCGCATCAGGATTTCACTACCATGAGGTCCACTG 248
QY 942 TAAACATATGACACGCTGGGAGCCGCTCTGCTGGGCGGTGGAATGTCTCCGCGG 1001
Db 249 CAACAACATGTTTGTAAAGGTGTGAAGATCTTTAAGGAGGTACAGTGTACTTCCGCTC 308
QY 1002 TGGCCCTCCATGGCTGGGAGGGGACCCATGCGCCACACACCTAGGGGCAACCTGGG 1061
Db 309 TG---AAGCAGTGTAGTGGGAACCTAATGCCATTTCTTCCCTTGTCTGGATGAGT 365
QY 1062 GGACCCAGAGCCCGGCTGTCTAGTCCCTTGGCGGCGGTGTGGCTCGCTTCTGCA 1121
Db 366 CAACCCAGTGTCTGTTTGTACCGGTGCTTCCACACCAAGATGGCCAGTGGCCATCAA 425
QY 1122 GTGCGCTTCTCTTGGGCGCCCTGTTACTTCTAGCGAATCTCTTCACTCTCTGA 1181
Db 426 GTGTCAATACCAATTTTGCAGATACCTGTGATGTTTCACTGAGATCATCTTCCCAATCAGA 485
QY 1182 TGTGTGTACAAATTCCTCTCCGGCACTGGGAGGACCTTCCCGCCAGCCCTCTGGTGGCC 1241

Db 486 TGTGCAATGTACAACAACCTCTGAAGCCCTGCCCCACCTCTCC----- 528
QY 1242 GCCTGCCACCTCCACCACCACTTCAGCAGCTTGGAGCTGGAGCCAGAGGCCAGCAGCC 1301
Db 528 -----TATGGCACCACCAACCTATGATCC 551
QY 1302 CGTGGCAAGCCGAGGGAGGCCGCCACCCCATCTCTATCGGCTGCTGGTGGCCATCAT 1361
Db 552 AATGCTTAAAGTTGATGACAGCAACTCGGATCTGATGGCTGCTGGTGGCCATCAT 611
QY 1362 CTTGCTCTCTGCTGCTATCATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1421
Db 612 CTTTATCTCTCTGCTGCTATCTGATCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 671
QY 1422 CAGCAAGGCTGAACGAGGCTGTTGAAGAGGAGCTGACGGTTCACCTCTCTGCTGCTGCT 1481
Db 672 GGAGAAGGCTTCTCGGAGGATGCTGATGATGAATGACATGACGCTTTCCTTCCGCAAG 731
QY 1482 GGACACTATCTCATCAACAACCGCCAGGCTCTAGAGAGCCACCCCGTACCAGGAGCC 1541
Db 732 TGATTTAGCATGTTCAACAATAACCGCTCTCTCACTCA----- 769
QY 1542 CCGGCTCTGCTGGAATCCGCCACTTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1601
Db 769 ---CCTAGTGAACAGGCTCCAACTCGACTTACGATCGCATCTTT----- 811
QY 1602 GCTCTCAATCAGCGCTACCGGCTCTCTTCTGCGCACTTACGCGCCCTGCGCCCTCGAGGCC 1661
Db 811 -----CCCTTCGCGCC 821
QY 1662 GGCCCCCCCCACACCGCTGGGCCAAACCCACCAACACCCAGGCTACAGTGGGGACTA 1721
Db 822 TGACTACAGGAGCCATCCAGGCTGATACGAACACTCCAGATTTGCTCCAGGGGAGA 881
QY 1722 TATGAGCTGTAGAGCCAGGCGCCGCTTCTGCGCCACCTCCCCAGAAACAGCGTCC 1781
Db 882 GGAGTCAGGCTGCGAGCGGTGTTGTGAAGCCAGTCCAGCCCTGAGGGGTGCG 941
QY 1782 CCATTATGCGAGGCTGACATTTACCTGCGAGGCGCTACCCGGGGCAACACCTATGC 1841
Db 942 CCACTATGAGAGGCTGACATAGTGAACCTCCAGAGGTGACAGGAGGCAACACATACTC 1001
QY 1842 TGTGCTCACTGCCCCCAGGGGCGAGTC---GGGATGGGCCCCCAGAGTGGATTCC 1898
Db 1002 AGTGCCTGCGTCACCATGCTGCTCTCAGSAAAGATGTGGCTGTGGAGAGTCC 1061
QY 1899 TCGATCTCACTCCGCTTCAAGGAGAGCTTGGCGAGGCGCAGTTTGGGAGGTGCACCT 1958
Db 1062 CAGGAAACTCTTAACCTTCAAGAGAGAGCTGGGAGAGGACAGTTTGGGAGGTTCATCT 1121
QY 1959 GTGTGAGGTGCGACAGCCCTCAAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2018
Db 1122 CTGTGAGTGGAGGAGTGAAGAAATTAAGAGAGAAAGATTTGGCCCTAGATGTCTGAGTGC 1181
QY 2019 GGGACACCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2078
Db 1182 CAACAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1241
QY 2079 GAATGATTTCTGAAAGAGGTGAAGATCATGTCGAGGCTCAAGAGCCCAACCAATCATTCG 2138
Db 1242 GAATGATTTCTTAAGAGAGATAAGATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1301
QY 2139 GCTGCTGCGGCTGTGTGTCAGGAGGACCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2198
Db 1302 TCTATGATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1361
QY 2199 CGCGACCTCAACAGTTCCTCAGTCCGCCACCCAGCTGGAGGACAGGAGCCGCGGGGCG 2258
Db 1362 TGGAGATCTCAATCAGTTCTTCTCCCGCCAGCGCCCTTAATTTCTCTCCAGCG--- 1418
QY 2259 CCCTGGGAGGCGAGGCTGCGCGAGGGGCCACCATCATGCTACCAATGCTGCTGCTGCTGCT 2318

Db 1418 -----ATGTACCGCACTGTCTAGTTACCACTGAAGTTTAT 1454
Oy 2319 GGCAGCCAGATCGCTCGGCGATGCGCTATCTGGCCACACTAACTTTGTATACATGGGA 2378
Db 1455 GGCATACCAAAATTCCTCTGGCATGAAGTACCTTCTCTCTAAATTTGTTCAACGGAGA 1514
Oy 2379 CTGTGGCCACCGGAAGTCTAGTTGGGGAATTTTCACCAACAAATCGCAGACTTTGG 2438
Db 1515 TCTGGCCACACGAACCTGTTTAGTGGTAAGAACTACAAATCAAGTACGTGACTTTGG 1574
Oy 2439 CATGAGCGGAACCTCTATCTGGGACTATTACCGTGTGCGGGCGCGGAGTGTGCGC 2498
Db 1575 AATGAGCAGAACTGTACAGTGTGACTATTACCGATCCAGGGCGCGGAGTGTGCGC 1634
Oy 2499 CATCCGCTGGATGCGCTGGGAGTGCATCTCATGGGAAAGTTACACGACTGCGAGTGA 2558
Db 1635 TATCCGCTGGATGCTTTGGGAGAGTATCTTGTGTGGCAAGTTTCACTACAGCAAGTGT 1694
Oy 2559 GTGGGCTTTGGTGTGACCTGTGGAGGTGCTGTCTGTAGGGCCGACGCTTTGG 2618
Db 1695 GTGGGCTTTGGGTTACTTTTGGGAGACTTTTACCTTTTGTCAAGAACAGCCCTATT 1754
Oy 2619 GCAGCTCACCGAGCAGAGTCTATGAGAACGCGGGGAGTTCTTCCGGGACCGAGGCG 2678
Db 1755 CCAGCTGTGAGTGAACAGAGTATTGAGAACTATGAGAGTTCTTCCGAGACCAAGGAG 1814
Oy 2679 GCAGGTGTACTGTCCGCGCGCTGCTGCGCGAGGCGCTATATGAGTGTATGCTGCG 2738
Db 1815 GCAGACTTACCTCCCTCAACCGACCTTTTGTCTGACTCTGTGTATAGCTGATGCTCAG 1874
Oy 2739 GTGCTGGAGCGGGAGTCTGAGCAGGACCGACCTTTTCCAGCTGCATCGGTTCT 2795
Db 1875 CTGCTGGAGAAGAGATACGAGAAGCCGCTCCTCATTCGAAGAAATCCACTTCTGCT 1931

RESULT 12

T22477
ID T22477 standard; cDNA to mRNA; 272 BP.
AC T22477;
DE Human gene signature HUMGS04088.
KW Gene signature; messenger RNA; mRNA: relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN WO9514772-A1.
PF 01-JUN-1995.
PR 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATSU) MATSUBARA K.
PA (OKUBU) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI: 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1: Page 1137; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (resp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.

SQ Sequence 272 BP; 52 A; 76 C; 72 G; 60 T;
Query Match 6.8%; Score 248; DB 1; Length 272;
Best Local Similarity 95.6%; Pred. No. 4.8e-42;
Matches 259; Conservative 0; Mismatches 11; Indels 1; Gaps 1;
Oy 2998 GATCCTCTCCACCTCTCTAGCATCCCTTGGGAGGAGTGGGGAGAAATATAGATAG 3057
Db 1 GATCCTCTCCACCTCTCTAGCATCCCTTGGGAGGAGTGGGGAGAAATATAGATAG 60
Oy 3058 ACATGGACATGGCCCTTGGAGCAGCTGGCCCTGACACACTGGACACTGATCTCTGGAGA 3117
Db 61 ACATGGACATGGCCCTTGGAGCAGCANNNGGCCCTGACACACTGGACACTGATCTCTGGAGA 120
Oy 3118 GGTGGCTGGCCCTGAGCTTCTCTCCCTGTACACACTGGACCTGGCTGGAGAAAT 3177
Db 121 GGTGGCTGGCCCTGAGCTTCTCTCCCTGTACACACTGGACACTGGCTGGAGAAAT 180
Oy 3178 CTGGGGGTGAGGAGGACAAGAGAGGAGAAATGTTTCCTTGTGCTGCTCTGTACTT 3237
Db 181 CTGGGGGTGAGGAGGACAAGAGAGGAGAAATNTTCTTGTGCTGCTCTGTACTT 240
Oy 3238 GTCCTCAGCTT-GGGCTTCTCTCTCTCTCCAT 3267
Db 241 GTCCTCAGCTTGGGGNTTCTCTCTCTCTCCAT 271

RESULT 13

T51456
ID T51456 standard; DNA; 2820 BP.
AC T51456;
DE 30-APR-1997 (first entry)
DE gD.trkA fusion useful in kinase receptor activation assay.
KW Kinase receptor activation; KIRA: receptor protein tyrosine kinase;
KW rPK: signal transduction; autophosphorylation; gD.trkA;
KW enzyme-linked immunosorbent assay; ELISA; ds.
OS Chimeric herpes simplex virus.
FH Key 5 Location/Qualifiers
FT misc_rna
FT /tag= a
FT /note= "sp6 RNA start"
FT misc_rna
FT /tag= b
FT /note= "begin gD from pchadIII"
FT cds
FT 293..2740
FT /tag= c
FT /product= gD.trkA fusion protein
FT signal_peptide
FT 380..460
FT /tag= d
FT /note= "trkA signal sequence"
FT mat_peptide
FT 461..2737
FT /tag= e
FT /note= "mature trkA"
FT poly_a_signal
FT 2778
FT /tag= f
FT /note= "SV40 early polyA"
FT WO9514930-A1.
PN 01-JUN-1995.
PD 18-NOV-1994; U13329.
PR 23-NOV-1993; US-157563.
PR 20-DEC-1993; US-170558.
PR 05-AUG-1994; US-286305.
PA (GETH) GENENTECH INC.
PI Godowski PJ, Mark MR, Sadick MD, Wong WLT;
DR WPI: 95-207042/27.
DR P-PSDB; W11940.
PT Measuring auto-phosphorylation of tyrosine kinase receptor - to
PT identify and characterise potential (ant)agonists of TKR
PS Example 3; Fig12a-g; 148pp; English.
CC DNA constructs (T51456-58) respectively code for gD.trkA (W11940),
CC gD.trkB (W11941) and gD.trkC (W11942), comprising fusions between

CC herpes simplex virus glycoprotein D flag polypeptide and human
CC receptor protein tyrosine kinases (rptks) trka, B and C, which are
CC involved in the interaction between neurotrophins and their target
CC cells. The gp.trk constructs can be expressed e.g. in dpl2.CHO
CC host cells and used in kinase receptor activation, enzyme-linked
CC immunosorbent assays (KIRA ELISA) to measure autophosphorylation of
CC rPTK kinase domains. These assays are used to identify and
CC characterise potential (ant)agonists of rPTKs, study ligand-receptor
CC interactions, measure phosphatase activity and identify rPTK ligands
CC in clinical samples.
SQ Sequence 2820 BP; 529 A; 867 C; 836 G; 588 T;

Query Match 5.4%; Score 195.2; DB 1; Length 2820;
Best Local Similarity 59.6%; Pred. No. 4.8e-31;
Matches 392; Conservative 0; Mismatches 248; Indels 18; Gaps 3;
QY 1988 TCAGTCTTGATTTCCCTTAAATGTCGTAAAGGACACCCCTTGTGCTAGCTGTCAAGA 2047
DB 1920 TCCTTCTGAGTGCCACACCTCTCTGCTGACGAGCAAGATGCTGTGCTGTCAAG- 1979
QY 2048 TCTTACGGCCAGATGCCACCAAGAAATGCCAGGAATGTTTCTTGAAGAGGTGAAGATCA 2107
DB 1979 --GCACCTGAAGGAGGCGTCCGAGAGTGTCTGGCAGGACTTCCAAGTGTGAGCTGAGCTGC 2036
QY 2108 TGTGAGGCTCAGGACCAACACATCATTCGGCTGCTGGGCTGTGTGTGAGGAGGACC 2167
DB 2037 TCACCATGCTGCAGCACCAGCAGCATCTGTCGCTTCTCGGCTGTGTCACCGAGGCGGCC 2096
QY 2168 CCCTCTGATGATTAATGACTACATGAGAGAGGAGGAGGAGCTCAACAGTTCTCTAGTGCC 2227
DB 2097 CCCTGCTCATGTCTTTGATGATATGCGCAGCGGAGCTCAACCGTCTCTCCGATCCC 2156
QY 2228 ACCAGCTGGAGGACAAAGCAGCCGAGGCGGCCCTTGGGAGCGGAGGCTGCGCAGGCGC 2287
DB 2157 ATGGACCTGATGCCAA-----GCTGCTGGCTGGTGGGAGGATGTGCTCCAG 2204
QY 2288 CCACCATCAGCTACCAATGCTGTCATGTGGCAGCCCGACAGATGCCCTCCGGCATGGCT 2347
DB 2205 GCCCCTGGGTCTGGGAGGCTGTGCGCGTGGCTAGCCAGGTGCTGCGGGGATGGTGT 2264
QY 2348 ATCTGGCCACACTCAACTTTGATACATCGGAGCTTGGCCACCGGAACCTGCTAGTGGGG 2407
DB 2265 ACCTGGCGGTCTGATTTGTGACCGGAGCTTGGCCACAGCAATGCTAGTGGGCC 2324
QY 2408 AAAATTTCAACATCAAAATCGCAGACTTTGGCATGAGCGGAACTCTATCTGGGACT 2467
DB 2325 AGGGACTGGTGTCAAGATTGGTATTTGGCATGAGCAGGAGATATCTACAGCACCAGCT 2384
QY 2468 ATTACCGTGTGACGGCGGCGAGTGTGCTCCATCCCTGGATGGCTGGGAGTGCATCC 2527
DB 2385 ATTACCGTGTGGGAGGCGCCACCATGCTGCCCCATTCGCTGATGCCGCCCGAGAGCATCC 2444
QY 2528 TCATGGGAAAGTTACAGACTCGAGTACGTGTGGGCTTTGGTGTGACCTGTGGGAGG 2587
DB 2445 TGTAACCGTAGTTCACACCGAGAGCGAGCTGTGGAGTTCGGCGTGGTGTCTGGGAGA 2504
QY 2588 TGCTGATGCTGTGAGGCCCGAGCCCTTTGGGAGCTACACGACGAGCAGTCAATCGA 2645
DB 2505 TCTTACCTACGGCAAG---CAGCCCTGGTACCAGCTCTCCAAACAGGAGGCAATCGA 2559

RESULT 14

ID V20445 standard; DNA; 2301 BP.
AC V20445;
DT 17-JUN-1998 (first entry)
DE Human c-trk oncogene.
KW Human; oncogene; proto-oncogene; neoplastic disease; anticancer;
OS Homo sapiens.
PN US5734039-A.
PD 31-MAR-1998.

PF 15-SEP-1994; 306691.
PR 15-SEP-1994; US-306691.
PA (UYJE-) UNIV JEFFERSON THOMAS.
PI Calabretta B. Skorski T;
DR WPI; 98-229882/20.
PT Anticancer composition comprising two anti-sense oligo:nucleotide(s)
PT - targeting cytoplasmic and nuclear oncogene(s)
PS Claim 1; Column 99-102; 92pp; English.
CC The present invention represents an oncogene from the present invention.
CC The present invention describes a composition which comprises two
CC antisense oligonucleotides. The first oligonucleotide is specific for a
CC cytoplasmic oncogene or proto-oncogene selected from ras, raf, Bcr-1,
CC c-fms, c-ros, c-kit, c-met, c-trk, c-src, c-abl, bcr-abl, c-fgr and
CC c-yes. The second oligonucleotide is specific for a nuclear oncogene or
CC proto-oncogene selected from myc, jun, c-ets, c-fos, c-myb, B-myb,
CC c-rel, c-vav, c-ski, c-spi, cyclin D1, PML/RAR alpha, AML1/MTG8,
CC E2A/p1 and ALL-1/AF-4. The composition is used for treating cancer.
CC The combination of antisense oligonucleotides has synergistically
CC enhanced ability to inhibit growth of cancer cells.
SQ Sequence 2301 BP; 545 A; 568 C; 785 G; 403 T;

Query Match 5.3%; Score 193.6; DB 1; Length 2301;
Best Local Similarity 59.4%; Pred. No. 9.6e-31;
Matches 391; Conservative 0; Mismatches 249; Indels 18; Gaps 3;
QY 1988 TCAGTCTTGATTTCCCTTAAATGTCGTAAAGGACACCCCTTGTGCTAGCTGTCAAGA 2047
DB 1275 TCCTTCTGAGTGCACACCTCTCTGCTGACGAGCAAGATGCTGTGCTGTCAAG- 1334
QY 2048 TCTTACGGCCAGATGCCACCAAGAAATGCCAGGAATGTTTCTTGAAGAGGTGAAGATCA 2107
DB 1334 --GCACCTGAAGGAGGCGTCCGAGAGTGTCTGGCAGGACTTCCAACGTGAGGCTGAGCTGC 1391
QY 2108 TGTGAGGCTCAGGACCAACACATCATTCGGCTGCTGGGCTGTGTGTGTCAGGAGGACC 2167
DB 1392 TCACCATGCTGACGACCAACACATCATTCGGCTGCTGGGCTGTGTGTGTCAGGAGGCGCC 1451
QY 2168 CCCTCTGTCATGATTACTGACTACTATGAGAGACGGGACCTCAACAGTTCCTCAGTGGCC 2227
DB 1452 CCCTGCTCATGCTTCTGAGTATATGCGCAGCGGAGGACCTCAACCGTCTCTCCGATCCC 1511
QY 2228 ACAGCTGGAGGACAAAGCAGCGGCGGCCCTTGGGAGCGGAGGCTGCGCAGGCGC 2287
DB 1512 ATGGACCCGATGCCAA-----GCTGCTGGCTGGTGGGAGGATGTGGCTCCAG 1559
QY 2288 CCACCATCAGCTACCAATGCTGTCATGTGGCAGCCCGACAGATCGCTCCGCGATGGCGT 2347
DB 1560 GCCCCTGGGTCTGGGCGAGCTGCTGGCGTGGCTAGCCAGTCTGCGGGGATGGTGT 1619
QY 2348 ATCTGGCCACACTCAACTTTGTACATCGGGAGCTGGCCACGCGGAACCTGCCCTAGTGGGG 2407
DB 1620 ACCTGGCGGTCTGCATTTTGTGACCGGAGCTGGCCACAGCAACTGCTTAGTGGGCC 1679
QY 2408 AAAATTTCAACATCAAAATCGCAGACTTTGGCATGAGCGCGGAACCTCTATGCTGGGACT 2467
DB 1680 AGGAGTGTGTGTAAGATTGGTATTTGGCATGAGCAGGAGATATCTACACACCGGACT 1739
QY 2468 ATTACCGTGTGAGGCGGCGGAGTGTGCTGCCCATCGCTGGATGGCCTGGGAGTGCATCC 2527
DB 1740 ATTACCGTGTGGAGGCGCGACCATGCTGCCCATTCGCTGATGCCGCCCGGAGAGCATCC 1799
QY 2528 TCATGGGAAAGTTACAGACTGCGAGTGTGGGCTTTGGTGTGAGCCCTGTGGGAGG 2587
DB 1800 TGTACCGTAGTTCACACCGAGAGGAGCTGTGAGCTTCGGCGTGGTGTCTCTGGGAGA 1859
QY 2588 TGCTGATGCTGTGTAGGGCCCGAGCCCTTTGGGAGCTCACCAGCAGGAGGAGTCAATCGA 2645
DB 1860 TCTTACCTACGGCAAG---CAGCCCTGGTACCAGCTCTCCAAACAGGAGGCAATCGA 1914

RESULT 15

Q99277

Result No.	Score	Query Match	Length	DB	ID	Description
1	4542	92.6	911	11	Q35407	Q35407 mus musculus
2	1154	23.5	220	11	Q64108	Q64108 rattus norv
3	968	13.7	791	5	Q19334	Q19334 caenorhabdi
4	924.5	18.9	183	11	Q64107	Q64107 rattus norv
5	702	14.3	700	5	Q18433	Q18433 geodia cydo
6	646	13.2	821	13	Q9YH44	Q9YH44 xenopus lae
7	646	13.2	811	13	Q9YH43	Q9YH43 xenopus lae
8	636.5	13.0	790	13	Q90699	Q90699 gallus gall
9	633.5	12.9	839	4	Q75682	Q75682 homo sapien
10	633.5	12.9	699	5	Q27656	Q27656 geodia cydo
11	628	12.8	503	4	Q15655	Q15655 homo sapien
12	627	12.8	591	4	Q15656	Q15656 homo sapien
13	625	12.7	422	4	Q43621	Q43621 homo sapien
14	621	12.7	282	13	Q91373	Q91373 xenopus lae
15	618	12.6	806	5	Q17576	Q17576 caenorhabdi
16	608.5	12.4	946	13	Q07153	Q07153 torpedo cal
17	606.5	12.4	868	11	Q62838	Q62838 rattus norv
18	603	12.3	868	11	Q61006	Q61006 mus musculus
19	603	12.3	871	11	Q61987	Q61987 mus musculus
20	602	12.3	737	5	Q18163	Q18163 caenorhabdi
21	601.5	12.3	869	4	Q15146	Q15146 homo sapien
22	601.5	12.3	881	11	Q61988	Q61988 mus musculus
23	600.5	12.2	860	11	Q61005	Q61005 mus musculus
24	591.5	12.1	685	5	Q24488	Q24488 drosophila
25	584	11.9	737	5	Q17305	Q17305 caenorhabdi
26	580	11.8	981	12	Q39477	Q39477 abelson mur
27	580	11.8	980	12	Q92809	Q92809 abelson mur
28	574	11.7	168	11	Q92209	Q92209 rattus norv
29	549.5	11.2	724	5	Q96391	Q96391 drosophila

Db 181 VELYGLWRDGLLSYAPVGTMTOLSEVMVHLNDSTVDYTAGGLQYGGGLQGLADGVVGL 240
Qy 239 DDFRKSQELRWPGYDYGVNSHNSFSGYVEMEFDFRLRAFAQMVHCHNMHTLGLARLP 298
Db 241 DDFRKSQELRWPGYDYGVNSHNSFSGYVEMEFDFRLRTFTQMCHNMHTLGLARLP 300
Qy 299 GGVCEFRFRRGPAWEGEPVHNSHNSFSGYVEMEFDFRLRAFAQMVHCHNMHTLGLARLP 358
Db 301 GGVCEFRFRRGPAWEGEPVHNSHNSFSGYVEMEFDFRLRAFAQMVHCHNMHTLGLARLP 360
Qy 359 FSEISFISDVVNSSPALGCTFPAPWMPGPPPTNFSSELSLEPRGQOPVAKEGSPTAI 418
Db 361 FSEISFISDVVNSSPALGCTFPAPWMPGPPPTNFSSELSLEPRGQOPVAKEGSPTAI 416
Qy 419 LIGCLVAIILLIILALMLRHLWRLLSKAERRVLEELTVHLSVPGDTILINNRPGP 478
Db 417 LIGCLVAIILLIILALMLRHLWRLLSKADSRVLEELTVHLSVPGDTILINNRPGP 476
Qy 479 REPPYQEPVPRGPPHSAPCVPGNSALLSNPAYRLLATYARPPGPPPTPAWAKPT 538
Db 477 REPPYQEPVPRGPPHSAPCVPGNSALLSNPAYRLLATYARPPGPPPTPAWAKPT 536
Qy 539 NTQAYSGDYMEPEKPGAPLPPPPQNSVPHYAEADIVTLQVGTGGNTYAYPALPPGAVGD 598
Db 537 NTQAYSGDYMEPEKPGAPLPPPPQNSVPHYAEADIVTLQVGTGGNTYAYPALPPGAVGD 596
Qy 599 GPPRVDFPRSLRKEKLGEGFCEVHCEVDSPQDLVSLDFPLNVRKHGPHLLVAVKILR 658
Db 597 GPPRVDFPRSLRKEKLGEGFCEVHCEVDSPQDLVSLDFPLNVRKHGPHLLVAVKILR 656
Qy 659 PDATKNARNDFLKVKIMSLRKLDPNIIIRLLGVCVQDDPLCMITDYMENGLNQLFSAHQL 718
Db 657 PDATKNARNDFLKVKIMSLRKLDPNIIIRLLGVCVQDDPLCMITDYMENGLNQLFSAHQL 716
Qy 719 EDKAEAGNPGQQAQGTISYPMLLHVAQIASGMRYLATLNFVHRDLATRNCLVGENF 778
Db 717 ENKATGLSGTSDQDGTISYPMLLHVAQIASGMRYLATLNFVHRDLATRNCLVGENF 776
Qy 779 TIKIADFGMSRLNAGDYRVVGRVAVLPIRWMAWECILMGKFTTASDVAFGVTLEVLML 838
Db 777 TIKIADFGMSRLNAGDYRVVGRVAVLPIRWMAWECILMGKFTTASDVAFGVTLEVLML 836
Qy 839 LCRAQPFGLTDEQVIENAGFEFFRQGRVYLSRPPACPGGLVLMRLCWSRSEQRPPF 898
Db 837 LCRAQPFGLTDEQVIENAGFEFFRQGRVYLSRPPACPGGLVLMRLCWSRSEQRPPF 896
Qy 899 SOLHRELAEDALNTV 913
Db 897 AQLHRELAEDALNTV 911
RESULT 2
ID Q64108 PRELIMINARY; PRT; 220 AA.
AC Q64108;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE RECEPTOR TYROSINE KINASE (FRAGMENT).
GN PTK-3L.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 95320273.
RA SAKUMA S., SAKA H., IJICHI A., TOFILON P.J.;
RT "Radiation induction of the receptor tyrosine kinase gene Ptk-3 in
normal rat astrocytes."
RL Radiat. Res. 143:1-7(1995).
DR EMBL; S77585; AAB34728.1; -.
ET NON_TER 1 1
SQ SEQUENCE 220 AA; 23670 MW; 807900DF CRC32;

Query Match 23.5%; Score 1154; DB 11; Length 220;
Best Local Similarity 98.2%; Pred. No. 8.5e-90;
Matches 216; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 404 GOOPVAKPEGSPATLIGCLVAIILLIILALMLRHLWRLLSKAERRVLEELTVHLS 463
Db 1 GOOPVAKPEGSPATLIGCLVAIILLIILALMLRHLWRLLSKAERRVLEELTVHLS 60
Qy 464 SVPGDTILINNRPGPREPPYQEPVPRGPPHSAPCVPGNSALLSNPAYRLLATYARP 523
Db 61 SVPGDTILINNRPGPREPPYQEPVPRGPPHSAPCVPGNSALLSNPAYRLLATYARP 120
Qy 524 PRGCPPTPAWAKPTNTQAYSGDYMEPEKPGAPLPPPPQNSVPHYAEADIVTLQVGTGG 583
Db 121 PRGCPPTPAWAKPTNTQAYSGDYMEPEKPGAPLPPPPQNSVPHYAEADIVTLQVGTGG 180
Qy 584 NTYAVPALPPGAVGDGPPRVDFPRSLRKEKLGEGFGE 623
Db 181 NTYAVPALPPGAVGDGPPRVDFPRSLRKEKLGEGFGE 220
RESULT 3
ID Q19354 PRELIMINARY; PRT; 791 AA.
AC Q19354;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE SIMILAR TO TYROSINE KINASE. NCBI GI: 1118035.
GN FL1D5.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 94150718.
RA WILSON R., AINSOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA PARDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
elegans."
RN [2]
RP SEQUENCE FROM N.A.
RA NHAN M.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA WATERSTON R.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U41532; AAB3276.1; -.
DR PFAM; PF00754; F5_F8_type-C; 1.
DR PFAM; PF00069; pkinase; 1.
SQ SEQUENCE 791 AA; 89990 MW; 991C9D67 CRC32;

Query Match 19.7%; Score 968; DB 5; Length 791;
Best Local Similarity 28.4%; Pred. No. 2.8e-73;
Matches 253; Conservative 152; Mismatches 309; Indels 178; Gaps 31;
Qy 46 DISASSSSSDSTAAR-HSRLESSDGDGAWCPAGSVFPEKEEYLVQDLQRLHLVALVGTQ 104
Db 34 DLNMLGKWFNGVFIRTAHQESGSGAWCPANQINSLSKEWLQISFSDVITVSVETQ 93

```

QY 105 RHAGLGKESRYRLYSRDG--RRWGWKDRWGQEVISGNEDEPGVVLKDLGPPMVARK 163
Db 94 RFDDGRMEYATFAKIOYWRPSLNAMASYKDDFELETIPANNDETEHAIRHDLRAIARR 153
QY 164 VREYPRADRVMSVCLRVLYGCLWRGLLSYTAPOGOTWYLSAVYLNDSYDGHVGL 223
Db 154 IRIVPVSNSTRVCMRVVEFGCFDLSVFNVDQGD---LQSGISYHDFSTDGNLANSF 210
QY 224 QY--GGGQLADGVVGLDDFKSKSELRLWPGYDYVGNHSNHSFGYVEMEFEDRLRAFOA 282
Db 211 HLTGGIGKLDYGEVGNKNNVFNH-----HKWGVNRK--RNGNVKLAFAFESLNRNIG 261
QY 283 MOVHCNNMHTLGLARPGVCECRFRPGPAMAWEGEPHRLGG-----NLGDP----- 331
Db 262 ILHTSN-----EFKKS-AKAFSSATVLFISNGKDFSDTIVHFNPNEDTES 306
QY 331 --ARAVSVPLGGVAREFLOCRFLFA--GPWLLFSEISFISDVVN-----NSSPALGCTPPP 382
Db 307 EVPRWIRIPVNNRIAKVAKIRLNFCTSDWLFISEVNFESNHTNIELLNDVVI----- 361
QY 383 APWPPGPPPTNFSSLELEPRGQOPVAKPEGSPAILIGCLVAIILLLLIALMLWRLH 442
Db 361 -----PDSVSYSFSTEHDGDT-----SMFAFII--FFFMFLIVAVIILTVLYRKR 403
QY 443 WRLLS-----KAERVLEEELTVHLSVPCDTILINNRPGPREPPYQPRGRNPHSA 497
Db 404 EYRKASSPNAKREIL-----LTIDGNTIKH-----VSPSTYOMAR----- 443
QY 498 PCVNGSALLSNPAYRLLIATYARPGPGPTPAWAKPTNT---QAYSGDYMEPEKPG 554
Db 443 ---DNLQNALIERKMPSPISDYAEPDISVCSDVTA-----NPLLYGIDGY-DTQKRS 493
QY 555 APLLPPPPQNSVPHYAEADIVTLQGVNTYAVPALPGAVGDPGRVDFPRSLRFEKE 614
Db 494 NPL-----SSMKYSYD-----GEVYCT-TLP-----EAIARDKLICVS 525
QY 615 KLCEGFGVHLCVDSPODLVSLDFPLNVRKGHPLLVAVKIILRPDATYKARNDFELVKV 674
Db 526 RICGGFGEVDLCOLEN-----RK-----VAVKKLH-GISOADEFSFHREIR 566
QY 675 IMSLKDPNIIRLLGVCVODDPLCMITDYMENGLDNOFLSAHQLEDKAAEGAPGDQAAO 734
Db 567 VLGSLLKHPNVVEVGVCTQKPLICIMEYENGDLKSYT-----LK 607
QY 735 GPTISVPMHLHVAQAQIAGMRYLATLNFVHRDLATRLNCLVGENFTIKIADFGMSRLYAG 794
Db 608 NPTIOTSQISICTLAAGLAVLESCNFVHRDIAARNCLVDGEGNVKYIADFGMARSLYSQ 667
QY 795 DYRVGQVRLPIRMWAWECILMGKETTASDYWAFQVTLWEVLMLCRAQPFQQLTDEQVI 854
Db 668 EYKVGKFLPIRMWAWALLGKFEASTASDVWGVGVTWMEIFSLCSEKPYSDMTDDVV 727
QY 855 ENAGEFFRDGQOVYLSRPPACQGLY-ELMLRCWSRESEQRPPFSQLHREL 905
Db 728 ENLQSMSTGSLKQVLSRPMCPKSLYNEQLPCWNYESSRPSFENVHLHL 779

RESULT 4
ID Q64107 PRELIMINARY; PRT; 183 AA.
AC Q64107;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
GN RECEPTOR TYROSINE KINASE (FRAGMENT).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RX SAKUMA S., SAYA H., IJICHI A., TOFILON P.J.;

```

```

RT "Radiation induction of the receptor tyrosine kinase gene Ptk-3 in
RL normal rat astrocytes."
RL Radiat. Res. 143:1-7(1995).
DR EMBL: S77556; AAB34729.1; -
FT NON_TER 1
SQ SEQUENCE 183 AA; 19717 MW; 0B1AD57B CRC32;

Query Match
Best Local Similarity 18.9%; Score 924.5; DB 11; Length 183;
Matches 179; Conservative 0; Mismatches 4; Indels 37; Gaps 1;

QY 404 GQPVAKPEGSPTAILGCLVAIILLLLIALMLWRLHRRLLSKAERRVLEELTVHL 463
Db 1 GQPVAKPEGSPTAILGCLVAIILLLLIALMLWRLHRRLLSKAERRVLEELTVHL 60
QY 464 SVPGDTILINNRPGPREPPYQPRGRNPHSPACVPNGSALLSNPAYRLLIATYARP 523
Db 61 SVPGDTILINNRPGPREPPYQPRGRNPHSPACVPNGSALLSNPAYRLLIATYARP 523
QY 524 PRGPGPTPAWAKPTNTQAYSGDYMEPEKPGAPLLPPPPQNSVPHYAEADIVTLQGV 583
Db 104 -----SGDYMEPEKPGAPLLPPPPQNSVPHYAEADIVTLQGV 143
QY 584 NTYAVPALPGAVGDPGRVDFPRSLRFEKEKLGEGQFGE 623
Db 144 NTYAVPALPGAVGDPGRVDFPRSLRFEKEKLGEGQFGE 183

RESULT 5
ID O18433 PRELIMINARY; PRT; 700 AA.
AC O18433;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DE 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE RECEPTOR TYROSINE KINASE.
OC Geodia cydonium (Sponge).
OC Eukaryota; Metazoa; Porifera; Demospongiae; Tetractinomorpha;
OC Astrophorida; Geodiidae; Geodia.
RN [1]
RP SEQUENCE FROM N.A.
RX GAMULIN V., SKOROKHOD A., MUELLER M., SCHAECKE H., MUELLER W.E.G.;
RT "Experimental indication in favor of the introns-late theory: the
RT receptor tyrosine kinase gene from the sponge Geodia cydonium."
EL J. Mol. Evol. 44:242-252(1997).
DR EMBL: X94128; CAA63848.1; -
DR PFAM; PF00047; Ig; 1.
DR PFAM; PF00069; pkinase; 1.
SQ SEQUENCE 700 AA; 77850 MW; 2F0C4179 CRC32;

Query Match
Best Local Similarity 14.3%; Score 702; DB 5; Length 700;
Matches 193; Conservative 85; Mismatches 191; Indels 120; Gaps 19;

QY 413 GSPTAILGCLVAIILLLLIAL--MLWRLHRR-----LISK 449
Db 140 GNSGVTAGVLIITLLLIALLIIFVFWVYRRKGFDLGSCRELSCGSCVPLAA 199
QY 450 AERVLEEELTVHLSVPGDTILINNRPGPREP-----PPYQPRGRN---PPHSAPCPVN 502
Db 200 LKGVKLPTRHENLNKNGTRLNERNHIADTTEIYSVQVKPLKKINKSPPLPP----- 256
QY 503 GSALLSNPAYRLLIATYARPGPGPTPAWAKPTNTQAYSGDYMEPEKPGAPLLPPP 562
Db 256 ---LTLFETELNELMSIDEKEELSPIOEKPT--RRNTGLSTYSQSGTIPKLAKLTKRKFK 311
QY 563 QNSVPHYAEADIVTLQ--GVTTGGNT--YAVPALP----- 593
Db 312 MKENPIQSDVILVLELELQVDNTLYALPLKPNSTRNSASFDTDDLASDPIYSVAINPSMF 371

```

```
QY 593 ---PGAVGD-----GP-----PRVDFPSRLRKEKLGSGQGEVHLCVD- 631
DB 372 TKRSITICNDLHPYGIYARPIKQKMRQPLNSVDNREVKGQGVAVLAETG 431
QY 631 -SPQDLVSLDFPLNVRKH-----PLLAVAKILRPDATKNARNDFLKEVKIMSLKDPNI 684
DB 432 LSGSNVASLP-----KGSMDAGVALVAVKLLKPDVSDVSRQSFDEIKFMSQLQHDSI 485
QY 685 IRLGVCVQDDPLCMITDYMENGLDNQFLSAHQLEDAAGAPDGGQAAGQPTISYPMLL 744
DB 486 VOLLAVCTHSHKPFVMEYMEENGLDNQFLQYQWV-----DSDALYSNQIPPTSTLL 537
QY 745 HVAQIAGSMRYLATLNFVHRDLATRNCLVGENFTIKIADFGMSRNLVAGDYRVQGRAV 804
DB 538 YMAQIAGSMRYLSSLTNVHRDLATRNCLVGENFTIKISDFGMSRNLVYRVYVRGRAM 597
QY 805 LPIRMAMECILMGKFTTASDVWAFVGLTLEVLMLCRAQPFQGLTDEQVIENAGEFFRDQ 864
DB 598 LPIRMAYTES-FYGRFSKSDAWAYGVTVWEIYTLGKKQPYEELDDQHMQDA---IRGT 653
QY 865 GRQVYLSRPPACPOGLYELMLRCWSRESEQRPPFSQHLRFLAEDALNTV 913
DB 654 GRRI-MGRPECPQAVYEVLLRCWEYAADRATFKEIH-----DSLNL 696

RESULT 6
QYH44 PRELIMINARY; PRT; 821 AA.
AC QYH44;
DT 01-MAY-1999 (TRENBLrel. 10, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)
DE NEUROTROPHIN RECEPTOR B XTRKB-ALPHA.
GN XTRKB.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Amphibia; Batrachia; Anura;
OC Mesobatrachia; Pipoidae; Pipidae; Xenopodinae; Xenopus.
RN [1]
RP SEQUENCE FROM N.A.
RA ISLAM N., GAGNON F., MOSS T.;
RT "Catalytic and non-catalytic forms of the neurotrophin receptor xtrkb
RT mRNA are expressed in a pseudo-segmental manner within the early
RT Xenopus central nervous system.";
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U39670; AAD00001.1;
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.
KW Receptor.
SQ SEQUENCE 821 AA; 92311 MW; DSCBCD27 CRC32;

Query Match 13.2%; Score 646; DB 13; Length 821;
Best Local Similarity 45.6%; Pred. No. 4.6e-46;
Matches 139; Conservative 48; Mismatches 80; Indels 38; Gaps 8;

QY 607 RSLRFEKELGSGQGEVHLC-----VDSQDVLSDPLNVRKHGHLPLVAVKILRPDATK 663
DB 534 RHNVILKRELGEAGFGKVFALCYNLYPEQD-----KILVAVKTLK-DASD 578
QY 664 NARNDFLEKVKIMSLRKDPNIIRLLGVCVQDDPLCMITDYMENGLDNQFLSAHQLEDAKAA 723
DB 579 NARKDFHREAELLTLNQHHEIVKFGVCEGDLIMVFYMKHGDNLKFLRAH----- 632
QY 724 EGAPGDQQAAG---PTISYPMLLHVAQAQIASGMRYLATLNFVHRDLATRNCLVGENFTI 780
DB 632 --GPDVLMAGENRPAELTQSOMLHIAQIAGWYLASQHFVHRDLATRNCLVGENLLV 689
QY 781 KIADFGMSRNLVAGDYRVQGRAVPIRMWAMECILMGKFTTASDVWAFVGLTLEVLMLC 840
DB 690 KIGDFGMSRDVSTDYRVGGHTMLPIRMWPPESIMYRKFTTESDVWSLGVVLWEIFTYG 749
QY 841 RAQPFQGLTDEQVIENAGEFFRDQGRQVYLSRPPACPOGLYELMLRCWSRESEQRPPFSQ 900
```

```
DB 750 K-QPWYQLSNNEVIECI-----TQGR--VLQRPRTCPKEVYDMLGCGWQREPHMRLNIKE 801
QY 901 LHRFL 905
DB 802 IHSLL 806

RESULT 7
QYH43 PRELIMINARY; PRT; 811 AA.
AC QYH43;
DT 01-MAY-1999 (TRENBLrel. 10, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)
DE NEUROTROPHIN RECEPTOR B XTRKB-ALPHA.
GN XTRKB.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Amphibia; Batrachia; Anura;
OC Mesobatrachia; Pipoidae; Pipidae; Xenopodinae; Xenopus.
RN [1]
RP SEQUENCE FROM N.A.
RA ISLAM N., GAGNON F., MOSS T.;
RT "Catalytic and non-catalytic forms of the neurotrophin receptor xtrkb
RT mRNA are expressed in a pseudo-segmental manner within the early
RT Xenopus central nervous system.";
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U39671; AAD00002.1;
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.
KW Receptor.
SQ SEQUENCE 811 AA; 91249 MW; B267BDE7 CRC32;

Query Match 13.2%; Score 646; DB 13; Length 811;
Best Local Similarity 45.6%; Pred. No. 4.5e-46;
Matches 139; Conservative 48; Mismatches 80; Indels 38; Gaps 8;

QY 607 RSLRFEKELGSGQGEVHLC-----VDSQDVLSDPLNVRKHGHLPLVAVKILRPDATK 663
DB 524 RHNVILKRELGEAGFGKVFALCYNLYPEQD-----KILVAVKTLK-DASD 568
QY 664 NARNDFLEKVKIMSLRKDPNIIRLLGVCVQDDPLCMITDYMENGLDNQFLSAHQLEDAKAA 723
DB 569 NARKDFHREAELLTLNQHHEIVKFGVCEGDLIMVFYMKHGDNLKFLRAH----- 622
QY 724 EGAPGDQQAAG---PTISYPMLLHVAQAQIASGMRYLATLNFVHRDLATRNCLVGENFTI 780
DB 622 --GPDVLMAGENRPAELTQSOMLHIAQIAGWYLASQHFVHRDLATRNCLVGENLLV 679
QY 781 KIADFGMSRNLVAGDYRVQGRAVPIRMWAMECILMGKFTTASDVWAFVGLTLEVLMLC 840
DB 680 KIGDFGMSRDVSTDYRVGGHTMLPIRMWPPESIMYRKFTTESDVWSLGVVLWEIFTYG 739
QY 841 RAQPFQGLTDEQVIENAGEFFRDQGRQVYLSRPPACPOGLYELMLRCWSRESEQRPPFSQ 900
DB 740 K-QPWYQLSNNEVIECI-----TQGR--VLQRPRTCPKEVYDMLGCGWQREPHMRLNIKE 791
QY 901 LHRFL 905
DB 792 IHSLL 796

RESULT 8
QYH43 PRELIMINARY; PRT; 790 AA.
AC QYH43;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)
DE TROPOMYOSIN RECEPTOR KINASE.
GN CTRKA.
```


OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria; Aves;
OC Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
RN [1]
RP SEQUENCE FROM N.A.
RA SCHROEPPEL A., V SCHACK D., DECHANT G., BARDE Y.A.;
RL Mol. Cell. Neurosci. 6:0-0(0).
DR EMBL; U43396; AAA85289.1; -.
DR PFAM; PF00560; LRR; 2.
DR PFAM; PF00069; pkinase; 1.
DR PROSITE; PS00107; PROTEIN KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN KINASE_TYR; 1.
DR PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.
SQ SEQUENCE 790 AA; 88669 MW; B5F7768C CRC32;

Query Match 13.0%; Score 636.5; DB 13; Length 790;
Best Local Similarity 30.8%; Pred. No. 2.7e-45;
Matches 192; Conservative 78; Mismatches 189; Indels 165; Gaps 22;

QY 356 WLFEISFTSDVNNSPALGGTFPPAPWPPPPPTNFSSLELEPRGQQPVAKPEGSP 415
DB 237 WELVLEINNISSLNK-----DLTCRAENAVGLAEDS- 270
QY 416 TAILGCLVAIIILLIILMLRLHWR-----RLSKAERRVLEELVHLVS- 466
DB 270 -VMLNVTFFPVILLSEAIPOHFWCIPFSDVSNPTPRILWFLFNGMLPEGYIHRIVEY 328
QY 466 -PGDTILN-----NRGPREPPYQPRGRGNPHPSAPCVNPS-ALLSNP---AYRL 516
DB 329 EPNSTVLHGCLQNR-----PTH-----VNNGTNLTVPVONPLGRATRSI 367
QY 517 LAYARPP--RGPGPTPAWAKPTNTQAYSGDYMEPEKPGAPLLPPPPONSYPHYAADI 574
DB 368 QGRFMDNPFSPPEEIPVISISPLGRNSS-----LEGVETADEHTFGVSV 414
QY 575 VTLQGV-----TCGNTYAYPA-----LPPGAVGGDP 600
DB 415 AVAVALFASLFLSVMLIALNKGHRKFGINRSVLAQEDDLAMSLHFMNLGSSPVST 474
QY 601 PRVD-----FPRSLRLEKEKLGEGOFGEVHLCEVDS--PDLYSL 638
DB 475 SKLDGLKSNFIENPOYECNACVHHVQRDIVLKWELGEGAFKGVLAECSHLLPQ----- 531
QY 639 DFPLNVRKGPLLVAVKILRPDATKNARNDFLKVKYKIMSLKADPNIIIRLLGVCVDDPLC 698
DB 531 -----EKLTVAVKALK-EVTESARLDQFREAELLTVLQHEHIVKFGVCTEGEPLI 580
QY 699 MITDYMENGLNOFLSAHOLEDKAAGAPGCGAAGQFTISYPMILLHVAQAISGRVLA 758
DB 581 MYFEYMKHGLNRLKSHGPDAKILD--OGOGQPCGQULTLSH--MLQIATQIASGMVYLA 636
QY 759 TLNFVHRDLATNCLVGENFTIKIADFGMSRLYAGDYRVQGRAVLPIRMAWECILMG 818
DB 637 SLHFVHRDLATNCLVGLDVLKVGDFGMSRDIYSTDYRVGGRVLMPIRMPPEILYR 696
QY 819 KETTASDVWAFGVTLEVMLCRAQPFQGLTDEQVIENAGEFFRDQGRQVYLSRPACQ 878
DB 697 KETTESDIWSFGVWLEIFTYKQ-OPWYQLSNTAEICI-----TQGRE--LERPRTCP 748
QY 879 GLVELMRCWSRESEORPFSOLH 902
DB 749 EYVDIMQSCWQREPOORIQDIH 772

RESULT 9
ID 075682 PRELIMINARY; PRT; 839 AA.
AC 075682;
DT 01-NOV-1998 (Tremblrel. 08, Created)
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
DT 01-MAY-1999 (Tremblrel. 10, Last annotation update)

DE TRKC PROTEIN.
GN TRKC.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Homnidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RA ICHASO N., RODRIGUEZ R., MARTIN-ZANCA D., GONZALEZ-SARMIENTO R.;
RL "Genomic characterization of the human TRKC gene";
DR EMBL; AJ224521; CAAL2029.1; JOINED.
DR EMBL; AJ224522; CAAL2029.1; JOINED.
DR EMBL; AJ224523; CAAL2029.1; JOINED.
DR EMBL; AJ224524; CAAL2029.1; JOINED.
DR EMBL; AJ224525; CAAL2029.1; JOINED.
DR EMBL; AJ224526; CAAL2029.1; JOINED.
DR EMBL; AJ224527; CAAL2029.1; JOINED.
DR EMBL; AJ224528; CAAL2029.1; JOINED.
DR EMBL; AJ224529; CAAL2029.1; JOINED.
DR EMBL; AJ224530; CAAL2029.1; JOINED.
DR EMBL; AJ224531; CAAL2029.1; JOINED.
DR EMBL; AJ224532; CAAL2029.1; JOINED.
DR EMBL; AJ224533; CAAL2029.1; JOINED.
DR EMBL; AJ224534; CAAL2029.1; JOINED.
DR EMBL; AJ224535; CAAL2029.1; JOINED.
DR PFAM; PF00560; LRR; 2.
DR PFAM; PF01463; LRRCT; 1.
DR PFAM; PF01462; LRRNT; 1.
DR PFAM; PF00089; pkinase; 1.
DR PROSITE; PS00107; PROTEIN KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN KINASE_TYR; 1.
DR PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.
SQ SEQUENCE 839 AA; 94428 MW; 986D7E71 CRC32;

Query Match 12.9%; Score 633.5; DB 4; Length 839;
Best Local Similarity 32.0%; Pred. No. 5.4e-45;
Matches 174; Conservative 66; Mismatches 132; Indels 171; Gaps 19;

QY 490 RGNPPHSPACVPNG-----SALLSNPA-----YRLLATYAR 522
DB 326 RGNPPHSLHWHNGQPLRESKIIHVEYQGEISEGCLLFNKRPTHYNNGYTLI-----AK 381
QY 523 PPRGPGPTPAWAKPTNTQAYSGDYME---PEKPGAPL---PPPPONSYPHYAADI 575
DB 382 NPLG-----TANQINGHFLKEPPESTDNFILFDEVPSTPIVTHKPEDIF 430
QY 576 TLOGVTGGNTYAVDAL-----P 592
DB 431 GVSIAVGLAAFAFACVLLVVLVFMINKYGRSKFGMKGPVAVISGEEDSASPLHHNHGITT 490
QY 593 PGAVGGDP-----PRVDFP-----RSRLRFEKELGEGQFG 622
DB 491 PSSLDAGPTVIGMTIRIPVIEPQVROGCHNCKPDTYVQHVKRRDVLKRELGEAGF 550
QY 623 EHLCEVDSQDLVSLDFPLNVRKGPLLVAVKILRPDATKNARNDFLKVKIMSLKDP 682
DB 551 KVFLAEC-----YNLSPTK-DKMLVAVKALK-DPTLAARKDFQREALLTNLQHE 598
QY 683 NITRLGVCVDDPLCMITDYMENGLNOFLSAHOLEDKAAGAPG-----DQAAAQGP- 737
DB 599 HIVFGVCGDGDPLIMVFEYMKHGLNKLFLAH-----GPDAMILLVDGQPRQAKG 649
QY 737 TISVPMILLHVAQAISGRVLAFLNFVHRDLATNCLVGENFTIKIADFGMSRLYAGDY 796
DB 650 ELGLSQMLHIASQIASGMVYLAHQHEVHRDLATNCLVGENFTIKIADFGMSRLYAGDY 709
QY 797 YR-----VOGRAVLPIRMAWECILMGKFTTASDVWAFGVTLEVMLCRA 842
DB 710 YRLFNPSGNDFCIWCVEVGGHTMLPIRMPPEISIMYKFTTESDVSFVILLHEITYGK- 769
QY 843 QPFGQLTDEQVIENAGEFFRDQGRQVYLSRPACQGLYELMRCWSRESEORPFSOLH 902
DB 843 QPFGQLTDEQVIENAGEFFRDQGRQVYLSRPACQGLYELMRCWSRESEORPFSOLH 902

Db 769 QPWFLSNTVEICI-----TQGR--VLERPRVCPKEVDYVLMGCMQREPQORLNKEIY 821

QY 903 RFL 905

Db 822 KIL 824

RESULT 10

Q27656 PRELIMINARY; PRT; 699 AA.

AC Q27656;

DT 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)

DE SPONGE RECEPTOR TYROSINE KINASE (EC 2.7.1.112)

DE (PROTEIN-TYROSINE KINASE) (TYROSYLPROTEIN KINASE) (PROTEIN KINASE

DE (TYROSINE)) (HYDROXYARYL-PROTEIN KINASE).

GN GCTK 2.

OS Geodia cydonium (Sponge).

OC Eukaryota; Metazoa; Porifera; Demospongiae; Tetractinomorpha;

OC Astrophorida; Geodiidae; Geodia.

RC [1]

RP SEQUENCE FROM N.A.

RX MEDLINE; 95142935.

RA GAMULIN V., RINKEVICH B., SCHAECKE H., KRUSE M., MUELLER I.M.,

RA WERNER E.G.;

RT "Cell adhesion receptors and nuclear receptors are highly conserved

RT from the lowest metazoa (marine sponges) to vertebrates.";

RL Biol. Chem. Hoppe-Seyler 375:583-588(1994).

RN [2]

RX SEQUENCE FROM N.A.

RP MEDLINE; 95251882.

RA SCHAECKE H., RINKEVICH B., GAMULIN V., MUELLER I.M., MUELLER W.E.G.;

RA "Immunoglobulin-like domain is present in the extracellular part of

RT the receptor tyrosine kinase from the marine sponge Geodia

RT cydonium.";

RL J. Mol. Biol. Evol. 44:242-252(1997).

CC -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP + PROTEIN

CC TYROSINE PHOSPHATE.

DR EMBL; X77528; CAA54663.1; -

DR PFAM; PF00047; 1g; 1.

DR PFAM; PF00069; pkinase; 1.

KW Transferase.

SQ SEQUENCE 699 AA; 77477 MW; F65E4B75 CRC32;

Query Match 12.9%; Score 633.5; DB 5; Length 699;

Best Local Similarity 32.0%; Pred. No. 4.1e-45;

Matches 194; Conservative 81; Mismatches 175; Indels 157; Gaps 25;

QY 413 GSTAILGLVAIIILLIILAL--MLWRLHWRLLSKAERRVLEELTVHLS----- 465

Db 140 GSNVGVIAGVLTLLIILIIILFVFWCYRR-----RGKLDLSCRELSGCGSC 193

QY 465 VPGDTILINRRPGPREPPYPRGRNPPHSPACVPNGSALLS-----NP 511

Db 194 VP---LLAALKCVKLPTRHRENLDR-----NGTRLRLNERNHIADTNTIYSV 238

QY 512 AYRELLATYARPRGCGPP-----TPAWAKPT--NT--QAYSGDY 547

Db 239 VQPLKAKISKSP--PLPPLTLTETELNETLSIDEKEULSQIQRTRRTGLSTISOG 296

QY 548 MEPEKPGAPLLPPPPQNSVPHYAEADIVTIQ-GVTGGNT-YAVPALP----- 593

Db 297 TIPKLAKLTKLRFKKNENPIYQSADELELELQVDNTLYALPSKPNSTRNSASFTDDL 356

QY 593 -----PGAVGD-----GP-----PRVDFPRSLRFEKEL 616

Db 357 ASDPIYSVAINPSMFTKRSSTIGNDLHPYGPYIARPIKQMRQPLNVSVDNIREVQI 416

QY 617 GEGQGEVHLCEVD--SPQDLYSLDFPLNVRKGH-----PLLVAVKILRPDANKARND 669

Db 417 GVGQFQGVAVLAETGLSGSNVASLP-----KGSMAAGVALVAVKKLKPQVSDVLSQF 470

QY 670 LKEVKIMSLKDPNIIRLIGVCVQDDPLCMITDYMENGDLNQFLSAHQLEDKAEGAPGD 729

Db 471 DKEIKFMSOLOHDSIVOLLAICTHSHKHPFIYMEYMGDLNQFLQKYQWVD-----D 522

QY 730 GQAAOGTTSYPMILHVAQAISGMRYLATLNFVHRDLATRNCLNGENFTKIADFGMSR 789

Db 523 DSALSNQNPSTLLYMAVQIASGMVYLSLNVYHRDLATRNCLVGSNFRKISDFGMR 582

QY 790 NLYAGDYRVQGRAVLPFRMMAWECILMGKFTASDVWAFVGTWLVEVLMLCRAQFPGOLT 849

Db 583 NLYERVYRVGRAMLPIRWMAES-FYGRFSEKSDAWAIGVTWVEIITLGGKQPYEELD 641

QY 850 DEQVIENAGEFFRDQGRQVYLSRP---PACQGLYELMLRCWSRESEQRPPESQLHRELA 906

Db 642 DQDMQDA---IRGTGRI-MGRPRGVACVRG---ATRCWVYAAADRATFKEIH---- 690

QY 907 EDALNTV 913

Db 690 -DSLNI 695

RESULT 11

Q13655 PRELIMINARY; PRT; 503 AA.

ID Q13655;

AC Q15655;

DT 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)

DE 55 KD PROTEIN.

GN TRK-T1.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

OC Eutheria; Primates; Catarrhini; Homiidae; Homo.

RP [1]

RN SEQUENCE FROM N.A.

RX MEDLINE; 92195650.

RA GRECO A., PIEROTTI M.A., BONGARZONE I., PAGLIARDINI S., LANZI C.,

RA DELLA PORTA G.;

RT "TRK-T1 is a novel oncogene formed by the fusion of TPR and TRK genes

RT in human papillary thyroid carcinomas.";

RL Oncogene 7:237-242(1992).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE; 92175499.

RA HILL K., BOONE C., GOEBL M., PUCCIA R., SDICU A.M., BUSSEY H.;

RT "Yeast KRE2 defines a new gene family encoding probable secretory

RT proteins, and is required for the correct N-glycosylation of

RT proteins.";

RL Genetics 130:273-283(1992).

DR EMBL; X62947; CAA44719.1; -

DR PFAM; PF00069; pkinase; 1.

DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.

DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.

DR PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.

SQ SEQUENCE 503 AA; 58174 MW; E2D051E1 CRC32;

Query Match 12.8%; Score 628; DB 4; Length 503;

Best Local Similarity 45.5%; Pred. No. 7.5e-45;

Matches 136; Conservative 49; Mismatches 84; Indels 30; Gaps 8;

QY 607 RSLRFRKEKLGEGQFGEVHLCEVDS---PDLYSLDFPLNVRKGHPLLVAVKILRPDANK 663

Db 214 RRDVLKWLKEGAFGKVFLEAECNLLPEQD-----KMLVAVKALK-EASE 258

```

QY 664 NARNDLKEVKIMSLKDPNIIRLLGVCVDDPLCMITDYMNGDNLQFSLAHLQEDKAA 723
Db 259 SARDFQREAEILTMLOHQHIVRFVCTEGRPLLMVFYMRHGDNLNRLSHGPDALL 318
QY 724 EGAPGCGQAAGPTISYPMMLHVAQAISGMRYLATLNFVHRDLATRNCLVGENETIKIA 783
Db 319 AGGE---DVAPGP-LGLGQLLAVASQAAGVYLAGLHFVHRDLATRNCLVGGGLVVKIG 374
QY 784 DFGMSRLNAGDYRVOGRAVLPPIRMMAWECILMGKFTTASDYWAFGVTLMWELMCLRAQ 843
Db 375 DFGMSRDYSTDYRVGGRMTLPIRMPPESILYRKFTTESDVWSEFGVYLWEIFYGK-Q 433
QY 844 PFGQLTDEQVIENAGEFFRDQGRVLSRPPACPGQIYELMLRCWSESEQRPPFSQLH 902
Db 434 PWQLSNTAIDCI-----TQGRE--LERPRACPEVYAIMRCWCQREPQQRHSIKDVH 485

RESULT 12
Q15656 PRELIMINARY; PRT; 591 AA.
AC Q15656;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DE 01-MAY-1999 (TReMBLrel. 10, Last annotation update)
OS P68 TRK-T3 ONCOPROTEIN.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RA GRECO A., MARIANI C., MIRANDA C., LUPAS A., PAGLIARDINI S., POMATI M.,
RX MEDLINE; 96025992.
RT novel gene on chromosome 3 whose product has a potential coiled-coil
RT domain.
RL Mol. Cell. Biol. 15:6118-6127(1995).
DR EMBL; X65960; CAA59936.1.
DR PFAM; PF00069; pkinase; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.
SQ SEQUENCE 591 AA; 65851 MW; 8DAAFAC4 CRC32;

```

```

Query Match 12.8%; Score 627; DB 4; Length 591;
Best Local Similarity 45.5%; Pred. No. 1.le-44;
Matches 136; Conservative 49; Mismatches 84; Indels 30; Gaps 8;

```

```

QY 607 RSLREKEKLGEGQFGEVHLCEVDS---PQDLVSLDFPLNVRKGPLLVAVKILRPDATK 663
Db 302 RRDIVLKEWELGEGAFGKVFLEASHNLLPEQD-----KMLVAVKALK-EASE 346
QY 664 NARNDLKEVKIMSLKDPNIIRLLGVCVDDPLCMITDYMNGDNLQFSLAHLQEDKAA 723
Db 347 SARQDFQREAEILTMLOHQHIVRFVCTEGRPLLMVFYMRHGDNLNRLSHGPDALL 406
QY 724 EGAPGCGQAAGPTISYPMMLHVAQAISGMRYLATLNFVHRDLATRNCLVGENETIKIA 783
Db 407 AGGE---DVAPGP-LGLGQLLAVASQAAGVYLAGLHFVHRDLATRNCLVGGGLVVKIG 462
QY 784 DFGMSRLNAGDYRVOGRAVLPPIRMMAWECILMGKFTTASDYWAFGVTLMWELMCLRAQ 843
Db 463 DFGMSRDYSTDYRVGGRMTLPIRMPPESILYRKFTTESDVWSEFGVYLWEIFYGK-Q 521
QY 844 PFGQLTDEQVIENAGEFFRDQGRVLSRPPACPGQIYELMLRCWSESEQRPPFSQLH 902
Db 522 PWQLSNTAIDCI-----TQGRE--LERPRACPEVYAIMRCWCQREPQQRHSIKDVH 573

RESULT 13
O43621 PRELIMINARY; PRT; 422 AA.
ID O43621

```

```

O43621;
AC 01-JUN-1998 (TReMBLrel. 06, Created)
DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
DE 01-MAY-1999 (TReMBLrel. 10, Last annotation update)
DE ETS RELATED PROTEIN-GROWTH FACTOR RECEPTOR TYROSINE KINASE FUSION
DE PROTEINS (FRAGMENT).
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RA KNEZEVICH S.R., MCPADDEN D.E., TAO W., LIM J.F., SORESENSEN P.H.B.;
RL Nat. Genet. 18:0-0(1998).
DR EMBL; AF041811; AAC12728.1.
DR PFAM; PF00069; pkinase; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.
FT NON_TER 1
SQ SEQUENCE 422 AA; 48156 MW; D6165427 CRC32;

Query Match 12.7%; Score 625; DB 4; Length 422;
Best Local Similarity 42.6%; Pred. No. 1e-44;
Matches 143; Conservative 51; Mismatches 90; Indels 52; Gaps 11;

QY 590 ALPGAVGDDGPPRVDFPRSLRKEKLGEGQFGEVHLCEVDSPODLVSLDFPLNVRKHP 649
Db 104 AMPIGRTADVQ---HIKRRDVLARELGEGAFGKVFLEAC-----YNLSPTK-DK 149
QY 650 LLVAVKILRPDATKNARNDLKEVKIMSLKDPNIIRLLGVCVDDPLCMITDYMNGDNL 709
Db 150 MLVAVKALK-DPTLAARKDQREAEILTMLOHQHIVRFVCTEGRPLLMVFYMRHGDNL 208
QY 710 NQFLSAHQLEDKAEGAPG-----DGOAAGP-TISTPMLLHVAQAISGMRYLATLNFV 763
Db 209 NKFLRAH-----GPDAMILVDGQPROAKGELGSLQMLHIASQIASGMVYLASOHV 259
QY 764 HRDLATRNCLVGENETIKIA DFGMSRLNAGDYR-----VQGRAVLPIRW 809
Db 260 HRDLATRNCLVGENETIKIA DFGMSRDYSTDYRLENPSGNDPFCIWCVEYGGHTMLPIRW 319
QY 810 MAWECILMGKFTTASDYWAFGVTLMWELMCLRAQFPQGLTDEQVIENAGEFFRDQGRVY 869
Db 320 MPPEISIMYRKFTTESDVWSEFGVYLWEIFYGK-QPWFQLSNTVEIECI-----TQGR--V 371
QY 870 LSRPPACPGQIYELMLRCWSESEQRPPFSQLHRL 905
Db 372 LERPRVCPKEVDYMLGCGWQREPQQRILNIKEIYKIL 407

RESULT 14
Q91373 PRELIMINARY; PRT; 282 AA.
ID Q91373
AC Q91373;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DE 01-MAY-1999 (TReMBLrel. 10, Last annotation update)
DE TRKB (FRAGMENT).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
OC Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae; Xenopodinae;
RN [1]
RP SEQUENCE FROM N.A.
RA COHEN-CORY S., FRASER S.E.;
RT "BDNF in the development of the visual system of Xenopus.";
RL Neuron 12:747-761(1994)
DR EMBL; S69713; AAB30791.1.
DR PFAM; PF00069; pkinase; 1.
DR PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.

```

FT NON_TER 1 1
SQ SEQUENCE 282 AA: 32820 MW: AA804CBA CRC32:

Query Match 12.78; Score 621; DB 13; Length 282;
Best Local Similarity 45.78; Pred. No. 1.3e-44;
Matches 132; Conservative 49; Mismatches 74; Indels 34; Gaps 8;

QY 607 RSRLREFKELGEGQGEVHLCVEVDSPQDLVSLDFPLNVRKGGHLLVAVKILRPDANKNA 665
DB 24 RINIVLKRELGEAGKGVFLAEC-----YNIYREQDKILVAVKTLK-DASDNA 70
QY 666 RNDFLKVKIMSRKLPNIIRLLGVCVODDPLCMITIDYMGNDLNLQFLSAHQLEDKAAEG 725
DB 71 RKDFHREAELLNLQHENIVKVFYGVCEGDPLIMVFYMKHGDNLKFLRAH----- 122
QY 726 APGDCQAAQGP---TISYPMLLHVAAQIATAGMRYLATLNFVHRDLATRNCLVGENFTIKI 782
DB 122 GPDAVMAEGNLLIELTOSQIMHISQQAAGMXYLASQHFVHRDLPTRNCLVGENLLVKI 181
QY 783 ADFGMSRLYAGDYRYVOGRAVLPIRWMAWECILMGKFTTASDVWAFVGLTWELMLCRA 842
DB 182 GDFGMSRDVSYDYRVGGMHMLPIRWMPESIMYRRFTTESDVWSGLVWLWEIFTYK- 241
QY 843 QPFGLTDEQVIENAGEFFRDQGRQVYLSRPPACPGQGLYELMLRCWSRE 891
DB 241 QPWYQLSNNEVIECI-----TQGR--VLQRPRTCPKEIYDLMRGQWORE 282

RESULT 15
ID Q17576 PRELIMINARY; PRT; 806 AA.
AC Q17576;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DE 01-MAY-1999 (TRENBLrel. 10, Last annotation update)
DE KIN-8 PROTEIN.
GN KIN-8.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
OC Rhabditina; Rhabditoidea; Rhabdientae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RA BERKS M.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 94150718.
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES M., KERSHAW J., KRISTEN J., LAISTER N., LATREILLE P.,
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA SMAILDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.,
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
DR EMBL; Z35595; CAA84639.1; -;
DR PFAM; PF00051; kring1e; 1.
DR PFAM; PF00069; pkinase; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
SQ SEQUENCE 806 AA: 90384 MW: 8505D993 CRC32:

Query Match 12.68; Score 618; DB 5; Length 806;
Best Local Similarity 36.28; Pred. No. 1e-43;
Matches 151; Conservative 55; Mismatches 151; Indels 60; Gaps 7;

QY 485 QBRPRGRNPPHSAPCVPNGSALLSNPAYRLLLATYARPPRGPPPTPAWAKPTWTQAYS 544
DB 375 QKTRHONAHCSAPSINSAA-----NSAYYRKLNGTSTPIMGRVPPHVE----- 420
QY 545 GYMEPEKEPCAPLLPPPPQNSVPHYAEADIVTLOGVTGGNTYAVPALPGAVGDGPPRVD 604
DB 420 ---MTSLPSAQHLGPPPPYPMDOHLOQARRPSPQEPIDNNSYKVFETIP----- 466
QY 605 FPRSRLRFKELGEGQGEVHLCVEVDSPQDLVSLDFPLNVRKGGHLLVAVKILRPDANKNA 664
DB 466 ---SOLSVREKIGEGQGVVH-----SGIYTSGLFAPEPMVAVKKCRHDATNA 511
QY 665 ARNDFLKEVKIMSRKLPNIIRLLGVCVODDPLCMITIDYMGNDLNLQFLSAHQLEDKAAE 724
DB 512 ERAQLEQETRAVATEDHPNVIKLVICVYMDNSLLAVFEYVMVHGDLHELLKVR----- 564
QY 725 GAPGDCQAAQGP---TISYPMLLHVAAQIATAGMRYLATLNFVHRDLATRNCLVGENFTIKI 784
DB 564 -VPPADHDMGGITEANAEFLYIATQALGMEYLASMSFVHRDLATRNCLVGDTRIAD 622
QY 785 FGMRSRLYAGDYRYVOGRAVLPIRWMAWECILMGKFTTASDVWAFVGLTWELMLCRAQP 844
DB 623 FGLMRTSYGSDYYKMLHRSWMPVRWMSKEAIEQGRFSEASDVWSFGVTLWEIWSFGR-QP 681
QY 845 FGQLTDEQVIENAGEFFRDQGRQVYLSRPPACPGQGLYELMLRCWSRESEQRPPFSQL 901
DB 682 YEGASNQQVIELY-----ANRHLLECPHNCPTNIYSLMVECWCHENIERRPTFSEI 731

Search completed: November 4, 1999, 01:53:54
Job time: 7614 sec

QY	1	MGPEALSSLLLLLVASGDADMKGHFPDPAKCRVALGHQMDRTPTDSDISASSWSNSTAAR	60
Db	1	MGPEALSSLLLLLVASGDADMKGHFPDPAKCRVALGHQMDRTPTDSDISASSWSNSTAAR	60
QY	61	HSRLSSDGDGAWCPAGSVFPKESEYLOVDLQRLHLVALVGTQGRHAGGLGKFEFSRYRL	120
Db	61	HSRLSSDGDGAWCPAGSVFPKESEYLOVDLQRLHLVALVGTQGRHAGGLGKFEFSRYRL	120

```

QY 121 RYSDGRWGWKDRGQEVISGNEDEGVVVKDLGPPMVARLVRYPRADRVMSVCLRV 180
Db 121 RYSDGRWGWKDRGQEVISGNEDEGVVVKDLGPPMVARLVRYPRADRVMSVCLRV 180
QY 181 ELYGCLWRDGLLSYTPVGTWYLSAVYNDSTYDGHVGVGLQGLQGLADGVVGLDD 240
Db 181 ELYGCLWRDGLLSYTPVGTWYLSAVYNDSTYDGHVGVGLQGLQGLADGVVGLDD 240
QY 241 FRKSOELRVWPGDYDYVWNSHSSGYSVEMEFEDRLAFQAMQVHCNNMHTLGARLPGG 300
Db 241 FRKSOELRVWPGDYDYVWNSHSSGYSVEMEFEDRLAFQAMQVHCNNMHTLGARLPGG 300
QY 301 VECFRFRGPAMAWEGEPMRHNLGNLGDPRARAVSVPLGGRVARFLQCRFLFAGPWLIFS 360
Db 301 VECFRFRGPAMAWEGEPMRHNLGNLGDPRARAVSVPLGGRVARFLQCRFLFAGPWLIFS 360
QY 361 EISFISDVVNSSPALGGTFPPAPWPPGPPPTNFSSLELEPRGQOPVAKPEGSPTAILI 420
Db 361 EISFISDVVNSSPALGGTFPPAPWPPGPPPTNFSSLELEPRGQOPVAKPEGSPTAILI 420
QY 421 GCLVAIILLLLIIALMLRHLWRRLLSKAERVRLEELTVHLSVPGDTILINNRPGPRE 480
Db 421 GCLVAIILLLLIIALMLRHLWRRLLSKAERVRLEELTVHLSVPGDTILINNRPGPRE 480
QY 481 PPPYQEPFRGNPPHSPACVPNGSALLSNPAYRLLLATYARPPRGPPPTPAWAKPTNT 540
Db 481 PPPYQEPFRGNPPHSPACVPNGSALLSNPAYRLLLATYARPPRGPPPTPAWAKPTNT 540
QY 541 QAYSGDYMEPEKPGAPLLPPPPQNSVPHYAEADIVTLOGVTGNTYAVPALPGAVGDGP 600
Db 541 QAYSGDYMEPEKPGAPLLPPPPQNSVPHYAEADIVTLOGVTGNTYAVPALPGAVGDGP 600
QY 601 PRVDFPRSRRLRFEKLGEGQFGEVHLCVDSQDLYSLDFFPLNVRKGHPLLVAVKILRPD 660
Db 601 PRVDFPRSRRLRFEKLGEGQFGEVHLCVDSQDLYSLDFFPLNVRKGHPLLVAVKILRPD 660
QY 661 ATKARNRDFLEKVKIMSRKDPNIIRLLGVCVQDDPLCMITDMYNGDNLQFSAHQLED 720
Db 661 ATKARNRDFLEKVKIMSRKDPNIIRLLGVCVQDDPLCMITDMYNGDNLQFSAHQLED 720
QY 721 KAEGAPCDGQAAGPTTISYPMLLHVAQAQASGMRYLATLNFVHRDLATRNCLVGENFTI 780
Db 721 KAEGAPCDGQAAGPTTISYPMLLHVAQAQASGMRYLATLNFVHRDLATRNCLVGENFTI 780
QY 781 KIADEGMSRLYAGDYRVQGRAVLPTRMMAWECILMGKFTTASDYWAFGYTLWEVLMCL 840
Db 781 KIADEGMSRLYAGDYRVQGRAVLPTRMMAWECILMGKFTTASDYWAFGYTLWEVLMCL 840
QY 841 RAQPFGLTDEQVIENAGEFFRQGRQVYLSRPPACQGLYELMLRCWSRESQRPFFSQ 900
Db 841 RAQPFGLTDEQVIENAGEFFRQGRQVYLSRPPACQGLYELMLRCWSRESQRPFFSQ 900
QY 901 LHRELAEDALNTV 913
Db 901 LHRELAEDALNTV 913

```

```

RESULT 2
US-08-336-343A-2
: Sequence 2, Application US/08336343A
: Patent No. 5677144
: GENERAL INFORMATION:
: APPLICANT: Ulrich, Axel
: APPLICANT: Alves, Frauke
: TITLE OF INVENTION: CKK-2, A No. 5677144el Receptor Tyrosine Kinase
: NUMBER OF SEQUENCES: 43
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Pennie & Edmonds
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: U.S.A.
: ZIP: 10036-2711

```

```

: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/336,343A
: FILING DATE: 08-NOV-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Coruzzi, Laura A.
: REGISTRATION NUMBER: 30,742
: REFERENCE/DOCKET NUMBER: 7683-065
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 790-9090
: TELEFAX: (212) 869-9741/8864
: TELEX: 66141 PENNIE
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 919 amino acids
: TYPE: amino acid
: TOPOLOGY: unknown
: MOLECULE TYPE: protein
: US-08-336-343A-2

```

```

Query Match 99.6%; Score 4882; DB 1; Length 919;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 912; Conservative 0; Mismatches 1; Indels 6; Gaps 1;

QY 1 MGPEALSSLLLLLVASGDADMGKHFDPKCRVALGMDRTIPDSISASSSSDSTAAR 60
Db 1 MGPEALSSLLLLLVASGDADMGKHFDPKCRVALGMDRTIPDSISASSSSDSTAAR 60
QY 61 HSRLESSDGDGAWCPAGSVPEKEEYLQVDLQRLHLVALVGTGHRAGGLGKFEFSYRL 120
Db 61 HSRLESSDGDGAWCPAGSVPEKEEYLQVDLQRLHLVALVGTGHRAGGLGKFEFSYRL 120
QY 121 RYSDGRWGWKDRGQEVISGNEDEGVVVKDLGPPMVARLVRYPRADRVMSVCLRV 180
Db 121 RYSDGRWGWKDRGQEVISGNEDEGVVVKDLGPPMVARLVRYPRADRVMSVCLRV 180
QY 181 ELYGCLWRDGLLSYTPVGTWYLSAVYNDSTYDGHVGVGLQGLQGLADGVVGLDD 240
Db 181 ELYGCLWRDGLLSYTPVGTWYLSAVYNDSTYDGHVGVGLQGLQGLADGVVGLDD 240
QY 241 FRKSOELRVWPGDYDYVWNSHSSGYSVEMEFEDRLAFQAMQVHCNNMHTLGARLPGG 300
Db 241 FRKSOELRVWPGDYDYVWNSHSSGYSVEMEFEDRLAFQAMQVHCNNMHTLGARLPGG 300
QY 301 VECFRFRGPAMAWEGEPMRHNLGNLGDPRARAVSVPLGGRVARFLQCRFLFAGPWLIFS 360
Db 301 VECFRFRGPAMAWEGEPMRHNLGNLGDPRARAVSVPLGGRVARFLQCRFLFAGPWLIFS 360
QY 361 EISFISDVVNSSPALGGTFPPAPWPPGPPPTNFSSLELEPRGQOPVAKPEGSPTAILI 420
Db 361 EISFISDVVNSSPALGGTFPPAPWPPGPPPTNFSSLELEPRGQOPVAKPEGSPTAILI 420
QY 421 GCLVAIILLLLIIALMLRHLWRRLLSKAERVRLEELTVHLSVPGDTILINNRPGPRE 480
Db 421 GCLVAIILLLLIIALMLRHLWRRLLSKAERVRLEELTVHLSVPGDTILINNRPGPRE 480
QY 481 PPPYQEPFRGNPPHSPACVPNGSALLSNPAYRLLLATYARPPRGPPPTPAWAKPTNT 540
Db 481 PPPYQEPFRGNPPHSPACVPNGSALLSNPAYRLLLATYARPPRGPPPTPAWAKPTNT 540
QY 541 QAYSGDYMEPEKPGAPLLPPPPQNSVPHYAEADIVTLOGVTGNTYAVPALPGAVGDGP 600
Db 541 QAYSGDYMEPEKPGAPLLPPPPQNSVPHYAEADIVTLOGVTGNTYAVPALPGAVGDGP 600
QY 601 PRVDFPRSRRLRFEKLGEGQFGEVHLCVDSQDLYSLDFFPLNVRKGHPLLVAVKILRPD 660
Db 601 PRVDFPRSRRLRFEKLGEGQFGEVHLCVDSQDLYSLDFFPLNVRKGHPLLVAVKILRPD 660

```

QY 661 ATKNA-----RNDLFKEVKIMSLKDPNIIRLLGVGVDDPLCMITDYMENGDLNQFLS 714
 Db 661 ATKNASFLSRNDFLKEVKIMSLKDPNIIRLLGVGVDDPLCMITDYMENGDLNQFLS 720
 QY 715 AHOLEDKAAGCAPDGOAAGPTISYPMLLHVAQAQASGMRYLATLNFVHRDLATRNCLV 774
 Db 721 AHOLEDKAAGCAPDGOAAGPTISYPMLLHVAQAQASGMRYLATLNFVHRDLATRNCLV 780
 QY 775 GENFTKIADFGMSRLYAGDYRVQGRAVLPPIRMMAWECILMKFTTASDWAFGVTLW 834
 Db 781 GENFTKIADFGMSRLYAGDYRVQGRAVLPPIRMMAWECILMKFTTASDWAFGVTLW 840
 QY 835 EVLMLCRAQPFQGLTDEQVIENAGEFFRDQGRQVYLSRPPACPGQGLYELMLRCWSRESQ 894
 Db 841 EVLMLCRAQPFQGLTDEQVIENAGEFFRDQGRQVYLSRPPACPGQGLYELMLRCWSRESQ 900
 QY 895 RPPFSQLHRFLAEDALNTV 913
 Db 901 RPPFSQLHRFLAEDALNTV 919

RESULT 3

US-08-336-343A-4
 ; Sequence 4, Application US/08336343A
 ; Patent No. 5677144
 ; GENERAL INFORMATION:
 ; APPLICANT: Ullrich, Axel
 ; APPLICANT: Alves, Frauke
 ; TITLE OF INVENTION: CCK-2, A No. 5677144el Receptor Tyrosine Kinase
 ; NUMBER OF SEQUENCES: 43
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10036-2711
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/336.343A
 ; FILING DATE: 08-NOV-1994
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Coruzzi, Laura A.
 ; REGISTRATION NUMBER: 30,742
 ; REFERENCE/DOCKET NUMBER: 7683-065
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 790-9090
 ; TELEFAX: (212) 869-9741/8864
 ; TELEX: 66141 PENNTE
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 855 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: protein
 ; US-08-336-343A-4

Query Match 49.3%; Score 2415; DB 1; Length 855;
 Best Local Similarity 52.1%; Pred. No. 7.3e-167;
 Matches 482; Conservative 118; Mismatches 227; Indels 98; Gaps 15;

QY 3 PRLSSLLLLLVAGSDMDKGFDPKCRYALGMQDRIPDSDISASSWSDSTAARHS 62
 Db 5 PRLMLVLLLLPILS---SAKAQVNPACRYPLGMSGGOIPDEDITASSQWSESTAAYG 61
 QY 63 RLESDGDGAWCPAGSVFPKE--EYLVQVLDLQRLHLVALVGTQGRHAGGLGKEFSRYLR 121

Db 62 RLDSEGDGAWCPPEIPVEDDLKEFLQIDLHTLHFLVLTQGRHAGGHIETAPMYKIN 121
 QY 122 YSRGRRWGHKDRWGEVIGSNGEDPEGVVLKDLGPPMVARLVRFYPRADRVMSVCLURVE 181
 Db 122 YSRDGTWISWRNRHKGQVLDGNSNPYDIFLKDLEPVARFVRFPVTDHSMNVCMRVE 181
 QY 182 LYGLWRDGLSLSTAPYQGMYL--SEAVYLNDSTYDGHVGGLOYGLQGLADGVVGLD 239
 Db 182 LYGCWLDGLVSNAPAGQQFVLPGGSIIILNDSVIDG--AVGYSMTGLGQLDGVSGLD 240
 QY 240 DFRKSQELRWPGDYDYVGSNNHSSFSGIVEMEEFEEDRLRAFOAMQVHCNNHHTLGARLP 299
 Db 241 DFTQTHEYHVWPGDYDYVGRNESATNGYIEIMFEEDIRNFTTMKVHCNNFAKGYKIFK 300
 QY 300 GVECFRRGRPMAGEGEMRHNLCGNLGDPRARAVSVPLGGRVARFLQCRFLFAGPWLFF 359
 Db 301 EVQCYF--RSEASEWEPNAISFFPLVLDVNP SARFVTVPLHHRMASAIKCOYHEADTWMF 359
 QY 360 SEISFISD--VYNNSSPALGGTFPPAPWPPGPPPTNFSSLELEPRGQOPVAKPEGSPAI 418
 Db 360 SEITFQSDAAMYNSEAL----PTSP-----MAPTYDPMKVDSDSNTRI 400
 QY 419 LIGCLVAIIILLLLIIALMLRHLRRLSKAERRVLEELTVHLSVPGDTIILNRR--P 476
 Db 401 LIGCLVAIIFILLAIIVILWRQFQWKMLEKASRRMLDDEMTVLSLSPSSSMFNRRSS 460
 QY 477 GPREP-----PPYQPRPRGNPPHSPACVPNGSALLSNPARYLLATYARP 523
 Db 461 SPSEQGSNSTYDRIFPLRPDYQEP-----SRLIRKLEP----- 495
 QY 524 PRGPGPTPAWAKPTQAYSGDYMEPEKPCAPLLPPPPQNSVPHVAEADIVTLQGVGG 593
 Db 495 -----APGEESGCGVWVPQSGP-----EGVPHAEADIVNLQGVGG 535
 QY 584 NTYAVPALPPGAVGDGPPRV--DFPRSLRFEKEKLGEGQFGEVHLCVDSQDLVSLDFPL 642
 Db 536 NTYSVPAVTMDLLSGKDVAVEEFPKLLTFKEKLGEGQFGEVHLCVDSQDLVSLDFPL 595
 QY 643 NVKRGHPLLVAVKILRPDATKNARNDFLEVKIMSLKDPNIIRLLGVGVDDPLCMITD 702
 Db 596 DVSANQVVLVAVKMLRADANKARNDFLEKIMSLKDPNIIRLLGVGVDDPLCMITE 655
 QY 703 YMENGDLNQFLSAHQLEDKAAAGAPGDGAAQOPTISYPMLLHVAQAQASGMRYLATLNF 762
 Db 656 YMENGDLNQFLSRHE-----PNSSSSDVTVSYTNLFKPMATQIASGMKYLSSNF 706
 QY 763 VHRDLATRNCLVGENFTIKIADFGMSRLYAGDYRVQGRAVLPPIRMMAWECILMGKFTT 822
 Db 707 VHRDLATRNCLVGNKTYTIKIDFGMSRLYSGDYRYRIQGRAVLPPIRMMSWESILLGKFTT 766
 QY 823 ASDVWAFGVTLWEVLMLCRAQPFQGLTDEQVIENAGEFFRDQGRQVYLSRPPACPGLYE 882
 Db 767 ASDVWAFGVTLWETFTTQEQPYSQLSDEQVIENAGEFFRDQGRQVYLSRPPACPGLYE 826
 QY 883 LMLRCWSRESEQRPPFSQLHRFLAE 907
 Db 827 LMLSCWRROTCKNRPSPQEIHLILLQ 851

RESULT 4

US-08-456-647B-20
 ; Sequence 20, Application US/08456647B
 ; Patent No. 5811516
 ; GENERAL INFORMATION:
 ; APPLICANT: Lemke Ph.D. et al., Greg E.
 ; TITLE OF INVENTION: PROTEIN-TYROSINE KINASE GENES
 ; NUMBER OF SEQUENCES: 54
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson P.C.
 ; STREET: 4225 Executive Square, Suite 1400
 ; CITY: La Jolla
 ; STATE: CA

COUNTRY: US
 ZIP: 92037
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/456,647B
 FILING DATE: 02-JUN-1995
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/237,401
 FILING DATE: 02-MAY-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/884,486
 FILING DATE: 15-MAY-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Wetherell Ph.D., John R.
 REGISTRATION NUMBER: 31,678
 REFERENCE/DOCKET NUMBER: 07251/007002
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 678-5070
 TELEFAX: (619) 678-5099
 INFORMATION FOR SEQ ID NO: 20:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 854 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-456-647B-20

Query Match 49.2%; Score 2413; DB 2; Length 854;

Best Local Similarity 52.3%; Pred. No. 1e-166;

Matches 481; Conservative 119; Mismatches 220; Indels 100; Gaps 15;

QY 9 LLLLLLVSAGDADMGKHFDPKACRYALQMDRTIPDSDISASSWSDSTAARHSLRSSD 68
 DB 10 VLLLLLLIUGSA--KAQNPATCRPLGMSGHIPDEDTASSQWSESTAAYKRLDSEE 67
 QY 69 GDGAWCPAGSVFPKE-EYLQVDLQRLHLVALVGTQGRHAGLGEKFSRSLRYSDGR 127
 DB 68 GDGAWCPETVPQDDLKLEFLQIDLRLFTLITLVTGQGRHAGHGEFAPMYKINYSRDS 127
 QY 128 RMWGWKDRGWQGVISGNDPEGVVLKLGPPMVARLVRYFPRADRVMSVCLRVLYGLW 187
 DB 128 RWISNRHKGKOVLDGNSNPYDFLKLDPPIVAREFVLIPVTHSMNVCMRVLYGCWV 187
 QY 188 RDGLLSYATAPVGTMYL--SEAVYINDSTYDGHVGLQGLQGLADGVVGLDDFRKSQ 245
 DB 188 LDGLVSYNAPAGQQFVLPGGSIIYLNDSYDG-AVGSYMTGLGLTQDGVSGLDLDTQTH 246
 QY 246 ELRVNPGYDYGWNSHSSSGVYVMEFEFDRLRAFOAMQVHCNNMHTLGARLPGVCECF 305
 DB 247 EYHVNPGYDYGWNSHSSSGVYVMEFEFDRIRNFTTKVHCNNMFAKGVKIFKEVQCYF 306
 QY 306 RRGPMAWGEPNRLNGLNGLPRARAVSVPLGGRVARFQCRFLFAGPWLLESEISFI 365
 DB 307 -RSEASENEPTAVYPLVLDVNPASRFTVPLHHRMASAKCOVHFADTWMPSEITFQ 365
 QY 366 SD--VYNNSPALGGTFFPAPWPPPTNFSLELEPRGQOPVAKPEGSGPTAILICGL 423
 DB 366 SDAAMYNNS--GALPTSP-----MAPTTYDPMKLVDDSDNTRILICGL 405
 QY 424 VAILLLLLLIALMLRWLRHLLSKAERVLLEELTVHLSVPGDTILINR---PGPR 479
 DB 406 VAIIFILLAIIVILWRQVQWLEKASRRMLDDEMTVSLSPSSSESMFNRRSSPSEQ 465
 QY 480 EP-----PPYQEPNPNPHSAPCVNGSALILSNPAYRLLLATYARPPRPG 528
 DB 466 ESNSTYDRIFPLRPDYQEP-----SRLIRKLPEF----- 495

QY 529 PPTPAWAKPTNTQAYSGDYMEPEKPGCAPLLPPPPQNSVPHYAEADIVTLQVGTGNTYAV 588
 DB 495 -----APGEEESCSCGVVKAQPNGP-----EGVPHYAEADIVNLQVGTGNTYCV 540
 QY 589 PALPGAVGDGPPRV-DPFRSLRFEKELGEGQFGEVHLCEVDSPODLVSLDFPLNVRKG 647
 DB 541 PAVTMDLISGKDVAVEEPKRLAFKELGEGQFGEVHLCEVGEKFKDKDFALDYSAN 600
 QY 648 HPLLVAVKILRPDATKNARNDFLKEVKIMSKDPNIIIRLLGVCVQDDPLCMITDYWENG 707
 DB 601 QPVLVAVKMLRADANKARNDFLKEIKIMSKDPNIIIRLLAVCITEDPLCMITEYMENG 660
 QY 708 DLNQFLSHOLEDKAAEGAPGQQAQPTTISYPMLLHVAQAQIASGMRYLATLNFVHRDL 767
 DB 661 DLNQFLSRHEPLSSCSDA-----TVSYANLKFMATQIASGMKYLSSLNFVHRDL 710
 QY 768 ATRNCLVGENFTIKIADFGMSRLYAGDYRVQGRAVLPIRMMAWECILMGKFTTASDVW 827
 DB 711 ATRNCLVGNKTYIKIADFGMSRLISGDIYRIQGRAVLPIRMMSHESILLGKFTTASDVW 770
 QY 828 AFGVTLWEVIMLCRAQFQGLTDEQVIENAGFEFFRDQGRQVYLSRPPACPOGLYELMLRC 887
 DB 771 AFGVTLWETFTFCQBPYSQLSDEQVIENAGFEFFRDQGRQVYLPALCPDSVYKMLSLC 830
 QY 888 WSRSEQRPPFSQHLRFLAE 907
 DB 831 WRRETKHRPESQEIHLQLQ 850

RESULT 5

US-08-237-401A-20
 ; Sequence 20, Application US/08237401A
 ; Patent No. 5837448
 ; GENERAL INFORMATION:
 ; APPLICANT: Lemke Ph.D. et al., Greg E.
 ; TITLE OF INVENTION: PROTEIN-TYROSINE KINASE GENES
 ; NUMBER OF SEQUENCES: 54
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson P.C.
 ; STREET: 4225 Executive Square, Suite 1400
 ; CITY: La Jolla
 ; STATE: CA
 ; COUNTRY: US
 ; ZIP: 92037
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/237,401A
 ; FILING DATE: 02-MAY-1994
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/884,486
 ; FILING DATE: 15-MAY-1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Haile Ph.D., Lisa A.
 ; REGISTRATION NUMBER: 38,347
 ; REFERENCE/DOCKET NUMBER: 07251/007001
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (619) 678-5070
 ; TELEFAX: (619) 678-5099
 ; INFORMATION FOR SEQ ID NO: 20:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 854 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-237-401A-20

Query Match 49.2%; Score 2413; DB 2; Length 854;

QY 379 TFPAPWPPGPPPTNFSSLEPRGQOPVAKPGSPTA 417
 Db 361 TFPAPWPPGPPPTNFSSLEPRGQOPVAKPGSPTA 399

RESULT 7
 US-08-701-191A-25
 ; Sequence 25, Application US/08701191A
 ; Patent No. 5942428

GENERAL INFORMATION:
 APPLICANT: Moosa Mohammadi, Joseph Schlessinger,
 APPLICANT: and Stevan R. Hubbard
 TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN
 TITLE OF INVENTION: OF NON-INSULIN RECEPTOR TYROSINE KINASE
 NUMBER OF SEQUENCES: 41
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Lyon & Lyon
 STREET: 633 West Fifth Street
 CITY: Suite 4700
 CITY: Los Angeles
 STATE: California
 COUNTRY: U.S.A.
 ZIP: 90071-2066
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 MEDIUM TYPE: storage
 MEDIUM TYPE: IBM Compatible
 OPERATING SYSTEM: IBM P.C. DOS 5.0
 SOFTWARE: Fastseq for Windows 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08701.191A
 FILING DATE: August 21, 1996
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Warburg, Richard J.
 REGISTRATION NUMBER: 32,327
 REFERENCE/DOCKET NUMBER: 227/088
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (213) 489-1600
 TELEFAX: (213) 955-0440
 TELEX: 67-3510
 INFORMATION FOR SEQ ID NO: 25:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 317 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-701-191A-25

Query Match 34.3%; Score 1680; DB 2; Length 317;
 Best Local Similarity 99.7%; Pred. No. 2.8e-114;
 Matches 316; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 597 GGGPPRVDFPRSLRFEKLGEGQFGEVHLCEVDSPQDLVSLDFPLNVRKGGHPLLVAVKI 656
 Db 1 GGGPPRVDFPRSLRFEKLGEGQFGEVHLCEVDSPQDLVSLDFPLNVRKGGHPLLVAVKI 60
 QY 657 LRPDATKNARNDFLKEVKIMSKDKPNIIRLLGVCVQDDPLCMITDYMGDNLQFLSAH 716
 Db 61 LRPDATKNARNDFLKEVKIMSKDKPNIIRLLGVCVQDDPLCMITDYMGDNLQFLSAH 120
 QY 717 QLEDKAAEGAPGQQAQGTTSYPMMLHVAQAISGMRYLATLNFVHRDLATRNCLVGE 776
 Db 121 QLEDKAAEGAPGQQAQGTTSYPMMLHVAQAISGMRYLATLNFVHRDLATRNCLVGE 180
 QY 777 NFTIKIADFGMSRLNLYAGDYRVQGRAVLPIRMAWECEILMGKFTTASDVWAFGVTLWEV 836
 Db 181 NFTIKIADFGMSRLNLYAGDYRVQGRAVLPIRMAWECEILMGKFTTASDVWAFGVTLWEV 240

QY 837 LMLCRAQPGQLTDEQVIENAGEFFRDQGRQVYLSRRPACPGQLYELMLRCWSESEQR 896
 Db 241 LMLCRAQPGQLTDEQVIENAGEFFRDQGRQVYLSRRPACPGQLYELMLRCWSESEQR 300
 QY 897 PFSQLHRFLAEDALNTV 913
 Db 301 PFSQLHRFLAEDALNTV 317

RESULT 8
 US-08-339-578-2
 ; Sequence 2, Application US/08339578
 ; Patent No. 5622862
 GENERAL INFORMATION:
 APPLICANT: Squinto, et al.
 TITLE OF INVENTION: ASSAY SYSTEMS FOR NEUROTROPHIN ACTIVITY
 NUMBER OF SEQUENCES: 2
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Regeneron Pharmaceuticals, Inc.
 STREET: 777 Old Saw Mill River Road
 CITY: Tarrytown
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10591-6707
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/339,578
 FILING DATE: 14-NOV-1994
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/690,199
 FILING DATE: 23-APR-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Kempler, Gail M.
 REGISTRATION NUMBER: 32,143
 REFERENCE/DOCKET NUMBER: 6526-061A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (914) 345-7400
 TELEFAX: (914) 345-7721
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 821 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-339-578-2

Query Match 13.5%; Score 659.5; DB 1; Length 821;
 Best Local Similarity 30.4%; Pred. No. 5.2e-40;
 Matches 201; Conservative 72; Mismatches 188; Indels 201; Gaps 22;
 QY 373 SPALGTFPPAPWPPGPPPTNFSSLEPRGQOPVAKPGSPTAILICGLVAILLL 432
 Db 217 SCVSGDPLTLYWDYGNLVSKHNETSHTQGLRITNISDDSGKQISCV 268
 QY 433 IIALMLRWRLHLLSKAERRVLEELTVHLSV---PGDTILINNRPGRPP 483
 Db 268 -----AENLVGEDQDSVNLTVHFAPTITFL-----ESPTSDHWC 302
 QY 483 -PYOEPRGRGNPHSPCPVNGS-----ALLLSNPA-----Y 513
 Db 303 IPP---TVRGNPDLQWFGNGAILNESKVICIKIHTNHTVHGCLQDLPNTHMNGDY 359
 QY 514 RLL-----ATYARPPRGPPPTPAWAKPTNTQAYSGDYMEP--E 551
 Db 360 TLMAKNEYKDKERQISAHEFGVGRVDYETNPVPEVLYEDWTPTDTI----GDTTKNSNE 415

```

QY 552 KPGAPLLPPPPQNSVPHYAEADIVTLQGV----- 581
Db 416 IPSTVDADQSNRHLVSIVAVVIVASVVGFCLLVWMLLLKLAHRSFGMKGPASVINSDD 475
QY 581 -----TGGNTYAVPALPGAVGDGP---PRVDFP-----RS 608
Db 476 SASPLHHISNGSNTPSSSEGGDAVITIGTKIPVIENPOYFGITNSQLKPDTFVQHKKR 535
QY 609 RLRFKEKLEGGEFGEVHLCE---VDSPODLVSLDPLNVRKRGHPLLVAVKILRPDATKNA 665
Db 536 NIVLKRELGEAGFKVFLAECYNLCPEQD-----KILVAVKTLK-DASDNA 580
QY 666 RNDFLKVKIMSLKDPNIIIRLLGVCVODDPLCMITDYMNGDNLNQLSAHQLED-KAAE 724
Db 581 RKDFHREALLNLOHEHIVKYGVCVEGDPLIMVFEYMKHGLKFLRAHGPDAVLMAE 640
QY 725 GAPGDQAAQAGPT-ISTPMLLHVAQAIAAGMRYLATLNFVHRDLATNCLVGENFTIKIA 783
Db 641 GNP-----PTLTQSOQLHIAQIAAGVYLASQHFVHRDLATNCLVGENLLVKIG 692
QY 784 DFGMSRLNLYAGDYRYVQGRVAVLPIRMWAEICILMGKFTTASDVWAFVTLWEVLMCLRAQ 843
Db 693 DFGMSRDVYSTDYRYVGGHTMLPIRMWPPESIMYRKFTTESDVWNSLGLVWEIFTYCK-Q 751
QY 844 PFGQLTDEQVIENAGEFFRDGQVYLSRPPACQGLYELMLRCWSRESORPPFSOLHR 903
Db 752 PWYQLSNNEVIECI-----TQGR--VLQRPRTCPQVIELMGLCQWREPHTRKNKSIHT 804
QY 904 FL 905
Db 805 LL 806

```

RESULT 9

```

US-08-469-537A-73
; Sequence 73, Application US/08469537A
; Patent No. 5843749
; GENERAL INFORMATION:
; APPLICANT: Maisonnier, et al.
; TITLE OF INVENTION: EHK AND ROR TYROSINE
; TITLE OF INVENTION: KINASES
; NUMBER OF SEQUENCES: 107
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Regeneron Pharmaceuticals, Inc.
; STREET: 777 Old Saw Mill River Road
; CITY: Tarrytown
; STATE: NY
; COUNTRY: U.S.A.
; ZIP: 10591
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,537A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USN 08/406,247
; FILING DATE: 17-MAR-1995
; APPLICATION NUMBER: USN 08/144,992
; FILING DATE: 28-OCT-1993
; APPLICATION NUMBER: USN 07/736,559
; FILING DATE: 26-JUL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Kempler, Ph.D., Gail M
; REGISTRATION NUMBER: 32,143
; REFERENCE/DOCKET NUMBER: REG 070C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 914-345-7400
; TELEFAX: 914-345-7721
; TELEX:

```

```

; INFORMATION FOR SEQ ID NO: 73:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 285 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-469-537A-73

```

```

Query Match 13.4%; Score 659; DB 2: Length 285;
Best Local Similarity 47.4%; Pred. No. 1.5e-40;
Matches 144; Conservative 45; Mismatches 79; Indels 36; Gaps 9;

```

```

QY 607 RSLRFKEKLEGGEFGEVHLCE---VDSPODLVSLDPLNVRKRGHPLLVAVKILRPDATK 663
Db 11 RHNIYVKRELGEAGFKVFLAECYNLCPEQD-----KILVAVKTLK-DASD 55
QY 664 NARNDFLKVKIMSLKDPNIIIRLLGVCVODDPLCMITDYMNGDNLNQLSAHQLED-KA 722
Db 56 NARKDFHREALLNLOHEHIVKYGVCVEGDPLIMVFEYMKHGLKFLRAHGPDAVL 115
QY 723 AEGAPGDQAAQAGPT-ISTPMLLHVAQAIAAGMRYLATLNFVHRDLATNCLVGENFTIK 781
Db 116 AEGNP-----PTLTQSOQLHIAQIAAGVYLASQHFVHRDLATNCLVGENLLVK 167
QY 782 IADFGMSRLNLYAGDYRYVQGRVAVLPIRMWAEICILMGKFTTASDVWAFVTLWEVLMCL 841
Db 168 IADFGMSRDVYSTDYRYVGGHTMLPIRMWPPESIMYRKFTTESDVWNSLGLVWEIFTYCK 227
QY 842 AQPFGQLTDEQVIENAGEFFRDGQVYLSRPPACQGLYELMLRCWSRESORPPFSOL 901
Db 228 -QPYQLSNNEVIECI-----TQGR--VLQRPRTCPQVIELMGLCQWREPHTRKNKNI 279
QY 902 HRFL 905
Db 280 HTLL 283

```

RESULT 10

```

US-08-286-305A-5
; Sequence 5, Application US/08286305A
; Patent No. 5766863
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J.
; APPLICANT: Mark, Melanie R.
; APPLICANT: Sadick, Michael D.
; APPLICANT: Shelton, David L.
; APPLICANT: Wong, Wai Lee Tan
; TITLE OF INVENTION: KINASE RECEPTOR ACTIVATION ASSAY
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/286,305A
; FILING DATE: 05-AUG-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/170558
; FILING DATE: 20-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/157563
; FILING DATE: 23-NOV-1993

```

ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 00,000
REFERENCE/DOCKET NUMBER: 854C1P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 847 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-286-305A-5

Query Match 13.4% Score 658; DB 2; Length 847;
Best Local Similarity 47.4% Pred. No. 6.9e-40;
Matches 144; Conservative 45; Mismatches 79; Indels 36; Gaps 9;

QY 607 RSLRFEKELGEGQGEVHLCE---VDSPODLVSLDFPLNVRKGHPLLVAVKILRPDANK 663
DB 560 RHNIYVKRELGEAGFGKVFLEACYNLCPEQD-----KILVAVKTLK-DASD 604

QY 664 NARNDLFKEVKIMSLKDPNIIRLLGVCVQDDPLCMITDYMENGLNOFLSAHQLED-KA 722
DB 605 NARKDFHREAELLTNLQHEHIVKFGVCVEGDPPLIMVFYMKHGLNKLRAHGPDAVLM 664

QY 723 AEGAPGDGQAAGQPT-ISYPMLLHVAQAISGMRYLATLNFVHRDLATRNCLVGENFTIK 781
DB 665 AEGNP-----PTELTQSOMLHIAQIAAGMVLASQHFVHRDLATRNCLVGENLLVK 716

QY 782 IADFGMSRLNLYAGDYRYVQGRAVLPIRNMWECILMGKFTTASDWAFGVTLWEVLMCLR 841
DB 717 IGDFGMSRDYSTDYRYVGGHTMLPIRNMPPESIMRYKFTTESDWLSGVVLEIFYGK 776

QY 842 AQPFGOLTDEQVIENAGEFFRDQGRQVYLSRPPACPGQLYELMLRCWSRESEQRPPFSOL 901
DB 777 -QPWYQLSNNVEIECI-----TQGR--VLQRPRTCPQEVYELMLGCGWQREPHMRKNIKGI 828

QY 902 HRFL 905
DB 829 HTLL 832

RESULT 11
US-08-359-705B-2
Sequence 2, Application US/08359705B
Patent No. 5844092
GENERAL INFORMATION:
APPLICANT: Presta, Leonard G.
APPLICANT: Shelton, David L.
TITLE OF INVENTION: Human trk Receptors and Neurotrophic Factor Inhibitors
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/359,705B
FILING DATE: 20-Dec-1994
CLASSIFICATION: 424
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/286845

FILING DATE: 08/10/94
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/215139
FILING DATE: 03/18/94
ATTORNEY/AGENT INFORMATION:
NAME: Torchia, PhD., Timothy E.
REGISTRATION NUMBER: 36,700
REFERENCE/DOCKET NUMBER: P0873P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-8674
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 822 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-359-705B-2

Query Match 13.4% Score 658; DB 2; Length 822;
Best Local Similarity 47.4% Pred. No. 6.6e-40;
Matches 144; Conservative 45; Mismatches 79; Indels 36; Gaps 9;

QY 607 RSLRFEKELGEGQGEVHLCE---VDSPODLVSLDFPLNVRKGHPLLVAVKILRPDANK 663
DB 535 RHNIYVKRELGEAGFGKVFLEACYNLCPEQD-----KILVAVKTLK-DASD 579

QY 664 NARNDLFKEVKIMSLKDPNIIRLLGVCVQDDPLCMITDYMENGLNOFLSAHQLED-KA 722
DB 580 NARKDFHREAELLTNLQHEHIVKFGVCVEGDPPLIMVFYMKHGLNKLRAHGPDAVLM 639

QY 723 AEGAPGDGQAAGQPT-ISYPMLLHVAQAISGMRYLATLNFVHRDLATRNCLVGENFTIK 781
DB 640 AEGNP-----PTELTQSOMLHIAQIAAGMVLASQHFVHRDLATRNCLVGENLLVK 691

QY 782 IADFGMSRLNLYAGDYRYVQGRAVLPIRNMWECILMGKFTTASDWAFGVTLWEVLMCLR 841
DB 692 IGDFGMSRDYSTDYRYVGGHTMLPIRNMPPESIMRYKFTTESDWLSGVVLEIFYGK 751

QY 842 AQPFGOLTDEQVIENAGEFFRDQGRQVYLSRPPACPGQLYELMLRCWSRESEQRPPFSOL 901
DB 752 -QPWYQLSNNVEIECI-----TQGR--VLQRPRTCPQEVYELMLGCGWQREPHMRKNIKGI 803

QY 902 HRFL 905
DB 804 HTLL 807

RESULT 12
US-08-286-846A-2
Sequence 2, Application US/08286846A
Patent No. 5877016
GENERAL INFORMATION:
APPLICANT: Presta, Leonard G.
APPLICANT: Shelton, David L.
TITLE OF INVENTION: Human trk Receptors and Neurotrophic Factor Inhibitors
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/286,846A
FILING DATE: 05-Aug-1994

```
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Torchia, PhD., Timothy E.
; REGISTRATION NUMBER: 36,700
; REFERENCE/DOCKET NUMBER: P0873P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-8674
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 822 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; US-08-286-846A-2

Query Match      13.4%; Score 658; DB 2; Length 822;
Best Local Similarity 47.4%; Pred. No. 6.6e-40;
Matches 144; Conservative 45; Mismatches 79; Indels 36; Gaps 9;

QY 607 RSLRFEKELGEGQGEVHLCE---VDSQDLVSLDFPLNVRKGHPLLVAVKILRPDATK 663
Db 535 RHNVILKRELGEAGFKVFLAECYNLCPQD-----KILVAVKTLK-DASD 579

QY 664 NARNDLKEVKIMSLKDPNIIRLLGVCVQDDPLCMITDYMENGDLNQLSAHQLED-KA 722
Db 580 NARKDFHREAELTNLQHEHIVKFKYGVCEGDPILMVFEYMKHGDNLKFLRAHGPDVILM 639

QY 723 AEGAPGDGQAAGQPT-ISYPMLLHVAQAQTAGMRYLATLNFVHRDLATNCLVGENFTIK 781
Db 640 AEGNP-----PTELTSQMLHIAQAQTAGMRYLASQHFVHRDLATNCLVGENLLVK 691

QY 782 IADFGMSRNLVAGDYRVQGRAVLPIRMWAVECILMGKFTTASDVWAFGVTLMVEVLMCR 841
Db 692 IGDGMSRDVYSTDYRVGGHTMLPIRWMPPEISIMYRKETTESDVWLSGLVWLWEIFTYCK 751

QY 842 AQPFGLTDEQVIENAGEFRDQGRQVYLSRPPACQGLYELMLRCWSRESQRPFFSOL 901
Db 752 -QPWYQLSNNEVIECI-----TQGR--VLQRPRTCPQEVYELMLGCGWQREPHMRKNKIGI 803

QY 902 HRFL 905
Db 804 HTLL 807

RESULT 13
US-08-441-104A-5
; Sequence 5, Application US/08441104A
; Patent No. 5891650
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J.
; APPLICANT: Mark, Melanie R.
; APPLICANT: Sadick, Michael D.
; APPLICANT: Shelton, David L.
; APPLICANT: Wong, Wai Lee Tan
; TITLE OF INVENTION: KINASE RECEPTOR ACTIVATION ASSAY
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 720 kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/441.104A
; FILING DATE: 15-MAY-1995
```

```
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/286305
; FILING DATE: 05-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/170558
; FILING DATE: 20-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/157563
; FILING DATE: 23-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 00,000
; REFERENCE/DOCKET NUMBER: 854C1PIC2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 847 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-441-104A-5

Query Match      13.4%; Score 658; DB 2; Length 847;
Best Local Similarity 47.4%; Pred. No. 6.9e-40;
Matches 144; Conservative 45; Mismatches 79; Indels 36; Gaps 9;

QY 607 RSLRFEKELGEGQGEVHLCE---VDSQDLVSLDFPLNVRKGHPLLVAVKILRPDATK 663
Db 560 RHNVILKRELGEAGFKVFLAECYNLCPQD-----KILVAVKTLK-DASD 604

QY 664 NARNDLKEVKIMSLKDPNIIRLLGVCVQDDPLCMITDYMENGDLNQLSAHQLED-KA 722
Db 605 NARKDFHREAELTNLQHEHIVKFKYGVCEGDPILMVFEYMKHGDNLKFLRAHGPDVILM 664

QY 723 AEGAPGDGQAAGQPT-ISYPMLLHVAQAQTAGMRYLATLNFVHRDLATNCLVGENFTIK 781
Db 665 AEGNP-----PTELTSQMLHIAQAQTAGMRYLASQHFVHRDLATNCLVGENLLVK 716

QY 782 IADFGMSRNLVAGDYRVQGRAVLPIRMWAVECILMGKFTTASDVWAFGVTLMVEVLMCR 841
Db 717 IGDGMSRDVYSTDYRVGGHTMLPIRWMPPEISIMYRKETTESDVWLSGLVWLWEIFTYCK 776

QY 842 AQPFGLTDEQVIENAGEFRDQGRQVYLSRPPACQGLYELMLRCWSRESQRPFFSOL 901
Db 777 -QPWYQLSNNEVIECI-----TQGR--VLQRPRTCPQEVYELMLGCGWQREPHMRKNKIGI 828

QY 902 HRFL 905
Db 829 HTLL 832

RESULT 14
US-08-457-880A-2
; Sequence 2, Application US/08457880A
; Patent No. 5910574
; GENERAL INFORMATION:
; APPLICANT: Leonard G. Presta
; APPLICANT: David L. Shelton
; APPLICANT: Roman Urfer
; TITLE OF INVENTION: HUMAN trk RECEPTORS AND NEUROTROPHIC FACTOR
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
```

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/457,880A
FILING DATE:
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/444,622
FILING DATE: 19-May-1995
APPLICATION NUMBER: 08/286846
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Torchia, PhD., Timothy E.
REGISTRATION NUMBER: 36,700
REFERENCE/DOCKET NUMBER: P0873PIC3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-8674
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 822 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-457-880A-2

Query Match 13.4%; Score 658; DB 2; Length 822;
Best Local Similarity 47.4%; Pred. No. 6.6e-40;
Matches 144; Conservative 45; Mismatches 79; Indels 36; Gaps 9;

Qy 607 RSLRFEKELGEGGEGVHLCE---VDSQDLVSLDFPLNVRKGHPDLLVAVKILRPD ATK 663
Db 535 RHNVILKRELGEAGFKVLAECYNLCPEQD-----KILVAVKTLK-DASD 579
Qy 664 NARNDFLKEVKIMSKDPNIIRLLGVCVQDDPLCMITDYMGNDLNOFLSAHQLED-KA 722
Db 580 NARKDFHREAEELLNLQHEHIVKFGVCVEGDPILMVFYMKHGLNKLFLRAHGPDAVLM 639
Qy 723 AEGAPGDQQAAGGT-ISTYPMLLHVAQAISGMRYLATLNFVHRDLATRNCLVGENFTIK 781
Db 640 AEGNP-----PTLQSQMLHQAQIAAGMVLASQHFVHRDLATRNCLVGENLLVK 691
Qy 782 IADFGMSRLNAGDYRYVQGRAVLPIRMWAECLMGKFTTASDVAFGVTLEVLMLCR 841
Db 692 IGDFGMSRDVSTDYRYVGGHTMLPIRMWPEIMYRKFTTESDVNSLGVVLEIFTYK 751
Qy 842 AQPFGQLTDEQVIENAGEFFRDQGRQVYLSRPPACPGQLYELMLRCWRSRESEQRPPFSOL 901
Db 752 -QPWYQLSNNEVIECI-----TQGR--VLPRTCPQEVYELMLGCGWQREPHMRKNIKGI 803
Qy 902 HRFL 905
Db 804 HTLL 807

RESULT 15
US-08-440-816A-5
Sequence 5, Application US/08440816A
Patent No. 5914237
GENERAL INFORMATION:
APPLICANT: Godowski, Paul J.
APPLICANT: Mark, Melanie R.
APPLICANT: Sadick, Michael D.
APPLICANT: Shelton, David L.
APPLICANT: Wong, Wai Lee Tan
TITLE OF INVENTION: KINASE RECEPTOR ACTIVATION ASSAY
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd

CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/440,816A
FILING DATE: 15-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/286305
FILING DATE: 05-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/170558
FILING DATE: 20-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/157563
FILING DATE: 23-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 00,000
REFERENCE/DOCKET NUMBER: 854CIP1C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 847 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-440-816A-5

Query Match 13.4%; Score 658; DB 2; Length 847;
Best Local Similarity 47.4%; Pred. No. 6.9e-40;
Matches 144; Conservative 45; Mismatches 79; Indels 36; Gaps 9;
Qy 607 RSLRFEKELGEGGEGVHLCE---VDSQDLVSLDFPLNVRKGHPDLLVAVKILRPD ATK 663
Db 560 RHNVILKRELGEAGFKVLAECYNLCPEQD-----KILVAVKTLK-DASD 604
Qy 664 NARNDFLKEVKIMSKDPNIIRLLGVCVQDDPLCMITDYMGNDLNOFLSAHQLED-KA 722
Db 605 NARKDFHREAEELLNLQHEHIVKFGVCVEGDPILMVFYMKHGLNKLFLRAHGPDAVLM 664
Qy 723 AEGAPGDQQAAGGT-ISTYPMLLHVAQAISGMRYLATLNFVHRDLATRNCLVGENFTIK 781
Db 665 AEGNP-----PTLQSQMLHQAQIAAGMVLASQHFVHRDLATRNCLVGENLLVK 716
Qy 782 IADFGMSRLNAGDYRYVQGRAVLPIRMWAECLMGKFTTASDVAFGVTLEVLMLCR 841
Db 717 IGDFGMSRDVSTDYRYVGGHTMLPIRMWPEIMYRKFTTESDVNSLGVVLEIFTYK 776
Qy 842 AQPFGQLTDEQVIENAGEFFRDQGRQVYLSRPPACPGQLYELMLRCWRSRESEQRPPFSOL 901
Db 777 -QPWYQLSNNEVIECI-----TQGR--VLPRTCPQEVYELMLGCGWQREPHMRKNIKGI 828
Qy 902 HRFL 905
Db 829 HTLL 832

Search completed: November 4, 1999, 01:52:00
Job time: 7590 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 1999, 13:36:38 ; Search time 22.34 Seconds
(without alignments)
968.013 Million cell updates/sec

Title: US-08-170-558-4
Perfect score: 4903
Sequence: 1 MGPEALSSLLLLLVASGDA.....QRPFSQHLRFALDAALNTV 913

Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues

Database: A_Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	4882	99.6	919	1	W34672 Human mammary carc
2	4875	99.4	919	1	W75502 Human mammary carc
3	4875	99.4	919	1	W75504 Human mammary carc
4	4852.5	99.0	914	1	R71100 Protein-tyrosine-k
5	4864.5	95.1	876	1	W34675 Human mammary carc
6	4651.5	94.9	882	1	W34673 Human mammary carc
7	3011	61.4	563	1	W34674 Human mammary carc
8	2415	49.3	855	1	W75503 Human colonic aden
9	2415	49.3	855	1	W75505 Human colonic aden
10	2415	49.3	855	1	W34671 CCK-2, a human mam
11	2415	49.3	855	1	W71114 Discoidin domain r
12	2413	49.2	854	1	W79152 Receptor protein t
13	2413	49.2	854	1	W81409 Receptor protein t
14	1825	37.2	650	1	R54089 Partial sequence o
15	658	13.4	822	1	R81630 Human trkB recepto
16	658	13.4	847	1	W11941 qD trkB fusion use
17	650.5	13.3	825	1	R81631 Human trkB recepto
18	650.5	13.3	850	1	W11942 qD trkB fusion use
19	642.5	13.1	825	1	R62021 Porcine trkB prote
20	639.5	13.0	825	1	R27148 Adult porcine trkB
21	639.5	13.0	825	1	R30883 trkB gene prod. Nu
22	638	13.0	814	1	W11940 qD trkB fusion use
23	637	13.0	830	1	R81627 Human trkB recepto
24	633.5	12.9	839	1	R81625 Human trkB recepto
25	627	12.8	728	1	R27149 Mouse trkB proto-o
26	627	12.8	728	1	R30884 Partial trkB gene
27	627	12.8	728	1	R62022 Murine trkB protei
28	625.5	12.8	839	1	W71621 Porcine trkB K2 is
29	621.5	12.7	739	1	R71618 Murine trkB K3 iso
30	606.5	12.4	868	1	W26507 Rat Dmk receptor.
31	606.5	12.4	868	1	W26610 Rat muscle-specifi
32	605.5	12.3	881	1	R84091 Nsk2 receptor with
33	603	12.3	868	1	R92717 Mouse muscle-local
34	603	12.3	871	1	W84087 Nsk2 receptor. Nuc
35	603	12.3	871	1	R62568 Mouse receptor tyr
36	603	12.3	867	1	W62583 Mouse receptor tyr
37	601.5	12.3	869	1	W26506 Human Dmk receptor
38	601.5	12.3	869	1	W26611 Human muscle-speci
39	601.5	12.3	861	1	W62572 Mouse Nsk2 (altern
40	601	12.3	873	1	R84092 Nsk2 receptor with
41	600.5	12.2	530	1	R92715 Mouse muscle-local
42	600.5	12.2	860	1	R92716 Mouse muscle-local
43	600.5	12.2	354	1	R94989 Nsk2 receptor intr

ALIGNMENTS

RESULT	ID	W34672	W34672 standard; Protein; 919 AA.	44	600.5	12.2	863	1	R84088	Nsk2 receptor with
	AC	W34672; 1998 (first entry)		45	600.5	12.2	863	1	W62569	Alternatively spli
	DE	Human mammary carcinoma kinase 10 (MCK-10) amino acid sequence.								
	KW	Mammary carcinoma kinase; MCK-10; receptor tyrosine kinase;								
	KW	proliferative disease; cancer; insulin receptor family;								
	KW	tyrosine kinase neurotrophin receptor; MCK-10 activity;								
	KW	neurological disorder; aberrant expression.								
	OS	Homo sapiens.								
	FT	Key	Location/Qualifiers							
	FT	Peptide	1..18							
	FT	Protein	/label= signal_sequence							
	FT	Domain	/note= "mature protein"							
	FT	Cleavage_site	31..185							
	FT		/label= Discoidin_I_like_domain							
	FT		304..307							
	FT		/label= endopeptidase_furin							
	FT		/note= "putative precursor cleavage site"							
	FT	Region	48..439							
	FT	Binding_site	/label= transmembrane_region							
	FT	Modified_site	617..627							
	FT		/label= ATP_binding_motif							
	FT		802..803							
	FT		/label= autophosphorylation_sites							
	FT		/note= "putative"							
	FT	Modified_site	798							
	FT		/label= autophosphorylation_site							
	FT		/note= "putative"							
	FT	Binding_site	844..847							
	FT		/label= binding_motif_for_P13_kinase							
	FT		/note= "binding motif for phosphatidylinositol 3' kinase"							
	FT	Binding_site	832..832							
	FT		/label= potential_substrate_binding_site							
	FT	Binding_site	506..509							
	FT		/label= putative_receptor_binding_site_for_SHC							
	FT		/note= "SHC is an oncogenic SH2 domain containing molecule"							
	FT	Binding_site	510..513							
	FT		/label= GTPase_activity_protein_binding_site							
	FT		/note= "putative"							
	FT	Region	505..541							
	FT		/note= "alternatively spliced variant"							
	FT	Region	666..671							
	FT		/note= "alternatively spliced sequence"							
	FT	Region	26..42							
	FT		/note= "antibody recognition sequence NTalpha"							
	FT	Region	309..321							
	FT		/note= "antibody recognition sequence NTbeta"							
	FT	Region	902..919							
	FT		/note= "antibody recognition sequence CTbeta"							
	PN	US5677144-A.								
	PD	14-OCT-1997.								
	PF	08-NOV-1994; 336343.								
	PR	16-NOV-1993; US-153397.								
	PA	(ALVE/) ALVES F H E.								
	PA	(ULLR/) ULLRICH A.								
	PI	Alves FHE, Ullrich A;								
	DR	WPI: 97-511869/47.								
	DR	N-FSDB; T93785.								
	PT	Truncated receptor tyrosine kinase CCK-2 - and nucleic acid coding								
	PT	for it, useful for cancer diagnosis								
	PS	Disclosure; Fig 1; 70pp; English.								

The present sequence represents the protein sequence of a mammary carcinoma kinase, called MCK-10. This kinase belongs to a novel family of receptor tyrosine kinases, and expression is associated with proliferative diseases such as cancer. The MCK-10 receptor tyrosine kinase has extensive sequence similarity to the insulin receptor family. The MCK-10 gene was isolated by PCR using 2 degenerate oligonucleotide primer pools, using a template cDNA synthesised by reverse transcription of poly-A RNA from the human mammary carcinoma cell line MCF7. The MCK-10 protein contains 2 alternative spliced sequences, from amino acids 505-541 and 666-671. The sequence represented by amino acids 585-595 may be important, as deletion of this motif in the activin receptor serine/threonine kinase results in reduced ligand binding affinity. MCK-10 is expressed in brain tissue, and the protein shares homology with the tyrosine kinase neurotrophin receptor. Modulation of MCK-10 activity therefore may be used for treatment of neurological disorders. MCK-10 is also expressed in a variety of cancer cell lines and tumour tissue. The nucleotide sequence of MCK-10, or parts of it, can be used for diagnostic purposes to detect aberrant expression of MCK-10 genes. Inhibitors of MCK-10 receptor activity may have therapeutic value in the treatment of diseases such as cancer.

Sequence 919 AA:

Query Match 99.6% Score 4882; DB 1; Length 919;
 Best Local Similarity 99.28; Pred. No. 0;
 Matches 912; Conservative 0; Mismatches 1; Indels 6; Gaps 1;

QY 1 MGPEALSLLLVSGDADMGKHFDPACRYALGMDQRTIPDSDISASSWSDDTAAR 60
 DB 1 MGPEALSLLLVSGDADMGKHFDPACRYALGMDQRTIPDSDISASSWSDDTAAR 60

QY 61 HSRLESSDGDGAWCPAGSVFPEEYQLVDLQRLHLVALVGTQGRHAGGLGKFFSYRL 120
 DB 61 HSRLESSDGDGAWCPAGSVFPEEYQLVDLQRLHLVALVGTQGRHAGGLGKFFSYRL 120

QY 121 RYSDGRWGWKDRWGQEVISGNEDEGVVYLDLGPMPVARLVFPRADRYMSVCLRV 180
 DB 121 RYSDGRWGWKDRWGQEVISGNEDEGVVYLDLGPMPVARLVFPRADRYMSVCLRV 180

QY 181 ELYGCLWRDGLSYTAPVQGTMYLSEAVYLNDSYDGTGHTVGGYGLGQLADGVVGLDD 240
 DB 181 ELYGCLWRDGLSYTAPVQGTMYLSEAVYLNDSYDGTGHTVGGYGLGQLADGVVGLDD 240

QY 241 FRKQELRVWPGYDYGVGNSHNSFSFGYVEMFEFEDRLRAFQAMQVHCNNHMTLGARLPGG 300
 DB 241 FRKQELRVWPGYDYGVGNSHNSFSFGYVEMFEFEDRLRAFQAMQVHCNNHMTLGARLPGG 300

QY 301 VECFRGCPAMAWEGEPNRLNGLNGLDPRARAVSVPLGGRVARFLQCRFLFAGPWLIFS 360
 DB 301 VECFRGCPAMAWEGEPNRLNGLNGLDPRARAVSVPLGGRVARFLQCRFLFAGPWLIFS 360

QY 361 EISFISDVVNNSSPALGGTFFPAPWPPGPPPTNFSSLELEPRGQQPVAKPEGSPTAILI 420
 DB 361 EISFISDVVNNSSPALGGTFFPAPWPPGPPPTNFSSLELEPRGQQPVAKPEGSPTAILI 420

QY 421 GCLVAIILLIILLIALLMLRWLRHLWRLSKARRVLEELIVHLSVPGDITLINNRGPRE 480
 DB 421 GCLVAIILLIILLIALLMLRWLRHLWRLSKARRVLEELIVHLSVPGDITLINNRGPRE 480

QY 481 PPPTQEPNPGNPHSPAPVNGSALLSNPAYRLLIATYARPPRGPPPTPAWAKPTNT 540
 DB 481 PPPTQEPNPGNPHSPAPVNGSALLSNPAYRLLIATYARPPRGPPPTPAWAKPTNT 540

QY 541 QAYSGDYMEPEKPCAPLLPPPPQNSVPHYAEADIVTLQGVGTGNTYAVPALPPGAVGDGP 600
 DB 541 QAYSGDYMEPEKPCAPLLPPPPQNSVPHYAEADIVTLQGVGTGNTYAVPALPPGAVGDGP 600

QY 601 PRVDFPSSRLREFKLEGGQFGEVHLCVDSPOQLVSLDFPLNVRKGHPLLVAVKILRPD 660
 DB 601 PRVDFPSSRLREFKLEGGQFGEVHLCVDSPOQLVSLDFPLNVRKGHPLLVAVKILRPD 660

QY 661 ATKNA-----RNDFLKEVKIMSRKDPNTIIRLLGVCVQDDPLCMITDYMENGDLNQELS 714
 DB 661 ATKNA-----RNDFLKEVKIMSRKDPNTIIRLLGVCVQDDPLCMITDYMENGDLNQELS 714

Db 661 ATKNASFSLFSRNDLKEVKIMSRKDPNTIIRLLGVCVQDDPLCMITDYMENGDLNQELS 720
 QY 715 AHQLEDKRAEGAPGQQAAGPTISYPMLLHVAAQIASGMRYLATLNFVHRDLATRNCLV 774
 Db 721 AHQLEDKRAEGAPGQQAAGPTISYPMLLHVAAQIASGMRYLATLNFVHRDLATRNCLV 780
 QY 775 GENFTIKIADFQMSNRLNLYAGDYRYVQGRAVLPIRMAWECILMGKFTTASDVWAFGVTLW 834
 Db 781 GENFTIKIADFQMSNRLNLYAGDYRYVQGRAVLPIRMAWECILMGKFTTASDVWAFGVTLW 840
 QY 835 EYVLMCLRAQPFQQLTDEQVIENAGEFFRDQGRQVYLSRPPACPOGLYELMLRCWSRESEQ 894
 Db 841 EYVLMCLRAQPFQQLTDEQVIENAGEFFRDQGRQVYLSRPPACPOGLYELMLRCWSRESEQ 900
 QY 895 RPPFQQLHRLFAEDALNTV 913
 Db 901 RPPFQQLHRLFAEDALNTV 919

RESULT 2
 R75502
 ID R75502 standard; Protein; 919 AA.
 AC R75502;
 DT 26-NOV-1995 (first entry)
 DE Human mammary carcinoma kinase 10 (MCK-10).
 KW Mammary carcinoma kinase 10; transmembrane receptor;
 KW receptor tyrosine kinase; cancer.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT peptide 1..18
 FT domain /label= signal
 FT /label= discoidin I-like domain
 FT cleavage_site 304..307
 FT region /label= putative precursor cleavage site
 FT /label= transmembrane
 FT misc_difference 505..541
 FT /label= alternatively spliced sequence I
 FT misc_difference 666..671
 FT /label= alternatively spliced sequence II
 FT misc_difference 25..42
 FT /label= NT alpha
 FT /note= "peptide antibody recognition site"
 FT misc_difference 309..321
 FT /label= NT beta
 FT /note= "see above"
 FT misc_difference 909..919
 FT /label= CT beta
 FT /note= "see above"
 PN W09514088-A.
 PD 26-MAY-1995.
 PF 16-NOV-1994; E03797.
 PR 16-NOV-1993; US-153397.
 PA (PLAC) MAX PLANCK GES. FORDERUNG WISSENSCHAFTEN.
 PI Alves FHE, Ullrich A;
 DR WPI; 95-224054/29.
 DR N-PSDB; Q92520.
 PT New nucleic acid encoding MCK-10 receptor tyrosine kinase - and
 PT derived vectors, transformed cells, proteins and antibodies useful
 PT for diagnosis and treatment of proliferative disease, esp. cancer,
 PT and for screening modulators
 PS Disclosure: Page 53-55; 115pp; English.
 CC cDNA prep. from human breast cancer cell line MCF7 (ATCC HTB22) and
 CC used in a PCR with two degenerate oligo primer pools based on
 CC conserved sequences of the kinase domain of receptor tyrosine
 CC kinases. One clone, designated MCK-10, was identified as novel RTK.
 CC The PCR fragment was used to screen a lambda gt11 library of human
 CC fetal brain cDNA. Several overlapping clones were identified. The
 CC composite of these cDNA clones is given in Q92520 and the deduced AA
 CC sequence in R75502. Some of the clones had a deletion of 6 AAs at
 CC posn. 2315 in the MCK-10 sequence. MCK-10 has all the
 CC characteristics of a receptor PK (see R75502 FT). Screening of

CC human placental library yielded two cDNA clones MCK-10-1 and
 CC MCK-10-2. One of the clones isolated from the human fetal brain
 CC library contd. an additional 18 nts in the TK domain. The MCK-10 splice
 CC isoforms have been designated MCK-10-1 (with an additional 111 bp between
 CC nts 1832 and 1943); MCK-10-2 (without any insertions); MCK-10-3 (with the
 CC additional 111 bp and 18 bp in the TK domain); and MCK-10-4 (with the
 CC additional 18 bp). The predicted mol. wts. of MCK-10-1 and MCK-10-2
 CC pre-receptors are 101.13 and 97.17 kD respectively, and can thus be
 CC subdivided into a 34.31 kD alpha subunit and a 66.84 or 62.88 kD
 CC beta subunits that contain the TK homology and alternative splice sites.
 SQ Sequence 919 AA;

Query Match 99.4%; Score 4875; DB 1; Length 919;
 Best Local Similarity 99.1%; Pred. NO. 0;
 Matches 911; Conservative 0; Mismatches 2; Indels 6; Gaps 1;

QY 1 MGPEALSSLLLLVVASGDADMKHFDPAKCRYALGMDRTTIPDSIDSSSSWSDSTAAR 60
 DB 1 MGPEALSSLLLLVVASGDADMKHFDPAKCRYALGMDRTTIPDSIDSSSSWSDSTAAR 60

QY 61 HSRLESSDGDGAWCPAGSVFPKEEYLVQDLQRLHLVALVGTQGRHAGGLGKEFSRSL 120
 DB 61 HSRLESSDGDGAWCPAGSVFPKEEYLVQDLQRLHLVALVGTQGRHAGGLGKEFSRSL 120

QY 121 RYSDGRRWGWKDRWGQEVISGNEDEPGVYLKDLGPPWVARLVRFYPADRVMSVCLRV 180
 DB 121 RYSDGRRWGWKDRWGQEVISGNEDEPGVYLKDLGPPWVARLVRFYPADRVMSVCLRV 180

QY 181 ELYGCLWRDGLLSYAPVQGMWYLSAAYLVNDSTVDGHTVGLGGLGGLADGVVGLDD 240
 DB 181 ELYGCLWRDGLLSYAPVQGMWYLSAAYLVNDSTVDGHTVGLGGLGGLADGVVGLDD 240

QY 241 FRKSOELRWPGYDYVGVNSHFSFGYVEMEFEDRLRAFAQMVHCCNMHTLGARLPGG 300
 DB 241 FRKSOELRWPGYDYVGVNSHFSFGYVEMEFEDRLRAFAQMVHCCNMHTLGARLPGG 300

QY 301 VECFRFRGPAMAWEGEPHRLNGLGNDPRARAVSVPLGGVAREFLQCFRLFAGPWLIFS 360
 DB 301 VECFRFRGPAMAWEGEPHRLNGLGNDPRARAVSVPLGGVAREFLQCFRLFAGPWLIFS 360

QY 361 EISFISDVVNNSSPALGGTFPPAPWPPGPPPTNFSSLELEPRGQOPVAKPEGSPPTALLI 420
 DB 361 EISFISDVVNNSSPALGGTFPPAPWPPGPPPTNFSSLELEPRGQOPVAKPEGSPPTALLI 420

QY 421 GCLVAIIILLLLIIALMLRWLHWRLLSKAERRVLEELTVHLSVPGDTILINNRPGPRE 480
 DB 421 GCLVAIIILLLLIIALMLRWLHWRLLSKAERRVLEELTVHLSVPGDTILINNRPGPRE 480

QY 481 PPYQEPFRGNPPHSAFCVNGSALLLSNPAYRLLLATYARPPRGPGPPTPAWAKPTNT 540
 DB 481 PPYQEPFRGNPPHSAFCVNGSALLLSNPAYRLLLATYARPPRGPGPPTPAWAKPTNT 540

QY 541 QAYSQDYMEPEKPGAPLLPPPPQNSVPHYAEADIVTLQGVGGNTYAVPALPPGAVGDGP 600
 DB 541 QAYSQDYMEPEKPGAPLLPPPPQNSVPHYAEADIVTLQGVGGNTYAVPALPPGAVGDGP 600

QY 601 PRVDFPRSLRFEKELGSGQGEVHLCVEVDSQDVLSDLDFPLNVRKGHPLLVAVKILRPD 660
 DB 601 PRVDFPRSLRFEKELGSGQGEVHLCVEVDSQDVLSDLDFPLNVRKGHPLLVAVKILRPD 660

QY 661 APTKNA-----RNDPLKVKIMSLKDPNIIIRLLGVCVQDDPLCMITDYMGDNLQFLS 714
 DB 661 APTKNAFSLFRNDPLKVKIMSLKDPNIIIRLLGVCVQDDPLCMITDYMGDNLQFLS 720

QY 715 AHQLEDKAAEGAPGDQAAGQPTTISYPMLLHVAQAQIASGMYLATLNFVHRDLATRNCLV 774
 DB 721 AHQLEDKAAEGAPGDQAAGQPTTISYPMLLHVAQAQIASGMYLATLNFVHRDLATRNCLV 780

QY 775 GENFIKTIADFGMSRLNLAGDYRYRQGRAVLPIRMAWECILMGKFTTASDVWAFGVTW 834
 DB 781 GENFIKTIADFGMSRLNLAGDYRYRQGRAVLPIRMAWECILMGKFTTASDVWAFGVTW 840

QY 835 EVLMCLRAQPFQGLTDEQVIENAGEFFRDGQVYLRRPACQGLYELMLRCWSESEQ 894
 DB 841 EVLMCLRAQPFQGLTDEQVIENAGEFFRDGQVYLRRPACQGLYELMLRCWSESEQ 900

QY 895 RPPFSQLHRLAEDALNTV 913
 DB 901 RPPFSQLHRLAEDALNTV 919

RESULT 3
 R75504
 ID R75504 standard; Protein; 919 AA.
 AC R75504;
 DT 26-NOV-1995 (first entry)
 DE Human mammary carcinoma kinase 10 (MCK-10).
 KW Mammary carcinoma kinase 10; MCK-10; transmembrane receptor;
 OS receptor tyrosine kinase; cancer.
 FH Homo sapiens.
 FT key
 FT peptide 1..18 Location/Qualifiers
 FT domain 31..185 /label= signal
 FT domain /label= discofoidin I-like domain
 FT cleavage_site 304..307 /label= discofoidin I-like domain
 FT region 417..439 /label= putative precursor cleavage site
 FT misc_difference 505..541 /label= transmembrane
 FT misc_difference 666..671 /label= alternatively spliced sequence I
 FT misc_difference 25..42 /label= alternatively spliced sequence II
 FT misc_difference 309..321 /label= NT alpha
 FT misc_difference 909..919 /note= "peptide antibody recognition site"
 FT misc_difference 909..919 /note= "see above"
 FT misc_difference 909..919 /label= CT beta
 FT misc_difference 909..919 /note= "see above"

WO9514089-A.
 26-MAY-1995.
 16-NOV-1994; E03799.
 16-NOV-1993; US-153397.
 (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
 PI Alves FHE, Ullrich A;
 WPI: 95-224055/29.
 N-PSDB: Q92522.

PT New nucleic acid encoding CCK-2 receptor tyrosine kinase - and
 PT derived vectors, transformed cells, proteins and antibodies, useful
 PT for diagnosis and treatment of proliferative and nervous system
 PT diseases and for screening modulators
 PS Disclosure; Page 70-72; 115pp; English.
 CC CDNA prepd. from human breast cancer cell line MCF7 (ATCC HTB22) was
 CC used in a PCR with two degenerate oligo primer pools based on
 CC conserved sequences of the kinase domain of receptor tyrosine
 CC kinases. One clone, designated MCK-10, was identified as novel RTK.
 CC The PCR fragment was used to screen a lambda g11 library of human
 CC fetal brain cDNA. Several overlapping clones were identified. The
 CC composite of these cDNA clones is given in Q92522 and the deduced AA
 CC sequence in R75504. Some of the clones had a deletion of 6AA at posn.
 CC 2315 in the MCK-10 sequence. MCK-10 has all the characteristics of
 CC a receptor PTK (see R75504 FT). Screening of human placental library
 CC yielded two cDNA clones. One of the clones isolated from the human
 CC fetal brain library contained an additional 18 nts in the TK
 CC domain. The MCK-10 splice isoforms have been designated MCK-10-1
 CC (with an additional 111 bp between nts 1832 and 1943); MCK-10-2
 CC (without any insertions); MCK-10-3 (with the additional 111 bps and
 CC 18 bp in the TK domain); and MCK-10-4 (with the additional 18 bp).
 CC The predicted mol. wts. of MCK-10-1 and MCK-10-2 pre-receptors are
 CC 101.13 and 97.17 kD respectively, and can thus be subdivided into a
 CC 34.31 kD alpha subunit and a 66.84 or 62.88 kD beta subunits that
 CC contain the TK homology and alternative splice sites.

SQ Sequence 919 AA;									
Query Match 99.4%; Score 4875; DB 1; Length 919;									
Best Local Similarity 99.1%; Pred. No. 0;									
Matches 911; Conservative 0; Mismatches 2; Indels 6; Gaps 1;									
Qy	1	MGPEALSSLLLLLVASGDADMKGHFPAKCRYALGMDRTIPDS	DISASSWS	STAAR	60				
Db	1	MGPEALSSLLLLLVASGDADMKGHFPAKCRYALGMDRTIPDS	DISASSWS	STAAR	60				
Qy	61	HSRLESSDGDGAWCPAGSVFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKFSRSYRL	120						
Db	61	HSRLESSDGDGAWCPAGSVFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKFSRSYRL	120						
Qy	121	HSRLESSDGDGAWCPAGSVFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKFSRSYRL	180						
Db	121	HSRLESSDGDGAWCPAGSVFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKFSRSYRL	180						
Qy	181	HSRLESSDGDGAWCPAGSVFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKFSRSYRL	240						
Db	181	HSRLESSDGDGAWCPAGSVFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKFSRSYRL	240						
Qy	241	HSRLESSDGDGAWCPAGSVFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKFSRSYRL	300						
Db	241	HSRLESSDGDGAWCPAGSVFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKFSRSYRL	300						
Qy	301	HSRLESSDGDGAWCPAGSVFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKFSRSYRL	360						
Db	301	HSRLESSDGDGAWCPAGSVFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKFSRSYRL	360						
Qy	361	HSRLESSDGDGAWCPAGSVFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKFSRSYRL	420						
Db	361	HSRLESSDGDGAWCPAGSVFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKFSRSYRL	420						
Qy	421	HSRLESSDGDGAWCPAGSVFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKFSRSYRL	480						
Db	421	HSRLESSDGDGAWCPAGSVFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKFSRSYRL	480						
Qy	481	HSRLESSDGDGAWCPAGSVFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKFSRSYRL	540						
Db	481	HSRLESSDGDGAWCPAGSVFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKFSRSYRL	540						
Qy	541	HSRLESSDGDGAWCPAGSVFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKFSRSYRL	600						
Db	541	HSRLESSDGDGAWCPAGSVFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKFSRSYRL	600						
Qy	601	HSRLESSDGDGAWCPAGSVFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKFSRSYRL	660						
Db	601	HSRLESSDGDGAWCPAGSVFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKFSRSYRL	660						
Qy	661	HSRLESSDGDGAWCPAGSVFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKFSRSYRL	714						
Db	661	HSRLESSDGDGAWCPAGSVFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKFSRSYRL	714						
Qy	715	HSRLESSDGDGAWCPAGSVFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKFSRSYRL	774						
Db	715	HSRLESSDGDGAWCPAGSVFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKFSRSYRL	774						
Qy	775	HSRLESSDGDGAWCPAGSVFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKFSRSYRL	834						
Db	775	HSRLESSDGDGAWCPAGSVFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKFSRSYRL	834						
Qy	835	HSRLESSDGDGAWCPAGSVFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKFSRSYRL	894						
Db	835	HSRLESSDGDGAWCPAGSVFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKFSRSYRL	894						
Qy	895	HSRLESSDGDGAWCPAGSVFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKFSRSYRL	913						
Db	895	HSRLESSDGDGAWCPAGSVFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKFSRSYRL	913						
Qy	901	HSRLESSDGDGAWCPAGSVFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKFSRSYRL	919						
Db	901	HSRLESSDGDGAWCPAGSVFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKFSRSYRL	919						
RESULT 4									

R71100									
ID R71100; Standard; Protein; 914 AA.									
AC R71100;									
DE 17-AUG-1995 (first entry)									
KW Protein-tyrosine-kinase PTK22.									
OS Homo sapiens.									
PN W09502187-A.									
PD 19-JAN-1995.									
PF 08-JUL-1994; G01480.									
PR 09-JUL-1993; GB-014271.									
PA (CANC-) CANCER RES INST.									
PI Barker KT, Crompton MR, Gusterson BA, Martindale JE;									
PI Mitchell PJ, Page MJ, Spence P;									
DR WPI: 95-066991/09.									
DR N-PSDB: Q84782.									
PT Method for screening substances, using protein tyrosine kinase -									
PT for potential utility as therapeutic agents for cancer									
PS Disclosure: Page 26-30; Sipp; English.									
CC cDNA derived from tumor metastatic tissue was amplified using									
CC primers (given in Q84783-84) based on sequences (R71101, R71103)									
CC associated with protein-tyrosine-kinases (PTK). Novel PTK22 was									
CC identified in an isolated subclone. The 3' sequence of PTK22 was									
CC obtained by reverse transcription (using the primer of Q84786) and									
CC PCR amplification (primers Q84787-88) of RNA of human breast									
CC carcinoma cell line MDA MG 468. The partial DNA sequence of PTK22									
CC is given in Q84782.									
SQ Sequence 914 AA;									
Query Match 99.0%; Score 4852.5; DB 1; Length 914;									
Best Local Similarity 99.2%; Pred. No. 0;									
Matches 907; Conservative 1; Mismatches 5; Indels 1; Gaps 1;									
Qy	1	MGPEALSSLLLLLVASGDADMKGHFPAKCRYALGMDRTIPDS	DISASSWS	STAAR	60				
Db	1	MGPEALSSLLLLLVASGDADMKGHFPAKCRYALGMDRTIPDS	DISASSWS	STAAR	60				
Qy	61	HSRLESSDGDGAWCPAGSVFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKFSRSYRL	120						
Db	61	HSRLESSDGDGAWCPAGSVFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKFSRSYRL	120						
Qy	121	HSRLESSDGDGAWCPAGSVFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKFSRSYRL	180						
Db	121	HSRLESSDGDGAWCPAGSVFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKFSRSYRL	180						
Qy	181	HSRLESSDGDGAWCPAGSVFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKFSRSYRL	240						
Db	181	HSRLESSDGDGAWCPAGSVFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKFSRSYRL	240						
Qy	241	HSRLESSDGDGAWCPAGSVFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKFSRSYRL	300						
Db	241	HSRLESSDGDGAWCPAGSVFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKFSRSYRL	300						
Qy	301	HSRLESSDGDGAWCPAGSVFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKFSRSYRL	360						
Db	301	HSRLESSDGDGAWCPAGSVFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKFSRSYRL	360						
Qy	361	HSRLESSDGDGAWCPAGSVFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKFSRSYRL	419						
Db	361	HSRLESSDGDGAWCPAGSVFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKFSRSYRL	419						
Qy	420	HSRLESSDGDGAWCPAGSVFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKFSRSYRL	479						
Db	420	HSRLESSDGDGAWCPAGSVFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKFSRSYRL	479						
Qy	480	HSRLESSDGDGAWCPAGSVFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKFSRSYRL	539						
Db	480	HSRLESSDGDGAWCPAGSVFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKFSRSYRL	539						
Qy	540	HSRLESSDGDGAWCPAGSVFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKFSRSYRL	599						
Db	540	HSRLESSDGDGAWCPAGSVFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKFSRSYRL	599						

```

Db 541 TQAYSGDYMEPRGAPLPPPPQNSVPHYAEADIVTLQGVGGNTYAVPALPPGAVGDG 600
QY 600 PPRVDFPRSLRKEKLGEGQFGEVHLCVEYDSDQDLVSLDFPLNVRKGHPLLVAVKILRP 659
Db 601 PPRVDFPRSLRKEKLGEGQFGEVHLCVEYDSDQDLVSLDFPLNVRKGHPLLVAVKILRP 660
QY 660 DATKNARNDLFKEVKTMSRLKDNRIIRLIGVCVQDDPLCMITDYMENGLDNFLSAHOLE 719
Db 661 DATKNARNDLFKEVKTMSRLKDNRIIRLIGVCVQDDPLCMITDYMENGLDNFLSAHOLE 720
QY 720 DKAAGAPGDGAQAQGTTSYPMLLHVAQAISGMRYLATLNFVHRDLATRNCLNGENFT 779
Db 721 DKAAGAPGDGAQAQGTTSYPMLLHVAQAISGMRYLATLNFVHRDLATRNCLNGENFT 780
QY 780 IKIADFGMSRLNLYAGDYRYVQGRAVLPIRMAWECILMGKFTTASDVMAFGVTLWEVLM 839
Db 781 IKIADFGMSRLNLYAGDYRYVQGRAVLPIRMAWECILMGKFTTASDVMAFGVTLWEVLM 840
QY 840 CRAQPGQLTDEQVIENAGEFFRDQGRQVYLSRPPACPOGLYELMLRCWSRESEORPPFS 899
Db 841 CRAQPGQLTDEQVIENAGEFFRDQGRQVYLSRPPACPOGLYELMLRCWSRESEORPPFS 900
QY 900 QLHRFLAEDALNTV 913
Db 901 QLHRFLAEDALNTV 914

RESULT 5
ID W34675 standard; Protein: 876 AA.
AC W34675;
DT 17-FEB-1998 (first entry)
DE Human mammary carcinoma kinase 10 (MCK-10) splice variant 3.
KW Mammary carcinoma kinase; MCK-10; receptor tyrosine kinase;
KW proliferative disease; cancer; insulin receptor family;
KW tyrosine kinase neurotrophin receptor; MCK-10 activity;
KW neurological disorder; aberrant expression.
OS Homo sapiens.
FH Key
FT 1. .18 Location/Qualifiers
FT Peptide
FT /label= signal_sequence
FT 19. .876
FT /note= "mature protein"
FT 31. .185
FT /label= Discoidin_I_like_domain
FT 304. .307
FT /label= endopeptidase_furin
FT /note= "putative precursor cleavage site"
FT 48. .439
FT /label= transmembrane_region
FT 580. .590
FT /label= ATP_binding_motif
FT 760. .761
FT /label= autophosphorylation_sites
FT /note= "putative"
FT 756. .756
FT /label= autophosphorylation_site
FT /note= "putative"
FT 802. .805
FT /label= binding_motif_for_pl3_kinase
FT /note= "binding motif for phosphatidylinositol 3' kinase"
FT 790
FT /label= potential_substrate_binding_site
FT 26. .42
FT /note= "antibody recognition sequence Ntalpha"
FT 309. .321
FT /note= "antibody recognition sequence Ntbeta"
FT 860. .877
FT /note= "antibody recognition sequence CTbeta"
FT US5677144-A.
FT 14-OCT-1997.

```

```

PF 08-NOV-1994; 336343.
PR 16-NOV-1993; US-153397.
PA (ALVE/) ALVES F H E.
PA (ULLR/) ULLRICH A.
PI Alves FHE, Ullrich A;
DR WPI; 97-511869/47.
PT Truncated receptor tyrosine kinase CKK-2 - and nucleic acid coding
PS for it, useful for cancer diagnosis
PS Disclosure; Page -: 70pp; English.
CC The present sequence represents a splice variant of a mammary
CC carcinoma kinase (MCK-10). This kinase belongs to a novel family
CC of receptor tyrosine kinases, and expression is associated with
CC proliferative diseases such as cancer. The MCK-10 receptor tyrosine
CC kinase has extensive sequence similarity to the insulin receptor family.
CC The MCK-10 gene was isolated by PCR using 2 degenerate oligonucleotide
CC primer pools, using a template cDNA synthesised by reverse transcription
CC of poly-A RNA from the human mammary carcinoma cell line MCF7. The
CC amplified PCR product was used to screen human foetal brain and placental
CC libraries, from which the present splice variant was isolated. This
CC splice variant does not possess amino acids 505-541 or 666-671 of MCK-10
CC (W34672). The sequence represented by amino acids 548-558 may be
CC important, as deletion of this motif in the activin receptor
CC serine/threonine kinase results in reduced ligand binding affinity.
CC MCK-10 is expressed in brain tissue, and the protein shares homology with
CC the tyrosine kinase neurotrophin receptor. Modulation of MCK-10 activity
CC therefore may be used for treatment of neurological disorders. MCK-10 is
CC also expressed in a variety of cancer cell lines and tumour tissue. The
CC nucleotide sequence of MCK-10, or parts of it, can be used for diagnostic
CC purposes to detect aberrant expression of MCK-10 genes. Inhibitors of
CC MCK-10 (or splice variants) receptor activity may have therapeutic value
CC in the treatment of diseases such as cancer.
CC note: the present sequence does not appear in the specification, but was
CC created using information provided.
SQ Sequence 876 AA;

```

```

Query Match 95.1%; Score 4664.5; DB 1; Length 876;
Best Local Similarity 95.8%; Pred. No. 0;
Matches 875; Conservative 0; Mismatches 1; Indels 37; Gaps 1;

QY 1 MGPEALSSLLLLLVASGDADMGKHPDPAKRYALGNQDRTIPDSISASSSSDSTAAR 60
Db 1 MGPEALSSLLLLLVASGDADMGKHPDPAKRYALGNQDRTIPDSISASSSSDSTAAR 60
QY 61 HSRLESSDGDGAWCPAGSVFPKEEYLQVDLQRLHLVALVGTQGRHAGGLGKFFSRYRL 120
Db 61 HSRLESSDGDGAWCPAGSVFPKEEYLQVDLQRLHLVALVGTQGRHAGGLGKFFSRYRL 120
QY 121 RYSDGRRMWGWKDRWGQEVISGNDEPGEVYLKDLGPPMVARLVRFYPRADRYMSVCLRV 180
Db 121 RYSDGRRMWGWKDRWGQEVISGNDEPGEVYLKDLGPPMVARLVRFYPRADRYMSVCLRV 180
QY 181 ELYGLWRDGLLSYTPVGTQMYLSEAVYLNDSTYDGHVTGGLOYGGLGADGVVGLDD 240
Db 181 ELYGLWRDGLLSYTPVGTQMYLSEAVYLNDSTYDGHVTGGLOYGGLGADGVVGLDD 240
QY 241 FRKSQELRVWPGDYDVGWSNHSFSSGYVEMEFEDRLRAFAQMVHCHNNHTLCARLPGG 300
Db 241 FRKSQELRVWPGDYDVGWSNHSFSSGYVEMEFEDRLRAFAQMVHCHNNHTLCARLPGG 300
QY 301 VECFRFRGPAMAWEGEPMRHNLGNLGDPRARAVSVPLGGRVARFLQCRFLFAGPMLLFS 360
Db 301 VECFRFRGPAMAWEGEPMRHNLGNLGDPRARAVSVPLGGRVARFLQCRFLFAGPMLLFS 360
QY 361 EISFISDVVNSSPALGCTFPAPMPPGPPPTNFSSELEPRGQQPVAKPESPTAILI 420
Db 361 EISFISDVVNSSPALGCTFPAPMPPGPPPTNFSSELEPRGQQPVAKPESPTAILI 420
QY 421 GCLVAITILLLLIIALMLRWRLHWRRLLSKAERRVLEELTVHLSPGDDTILINRPGPRE 480
Db 421 GCLVAITILLLLIIALMLRWRLHWRRLLSKAERRVLEELTVHLSPGDDTILINRPGPRE 480
QY 481 PPYQEPFRPNPNPHSAPCPVNGSALLLSNPAYRLLLATYARPPRPGPPPTPAWAKPTNT 540

```

Db 481 PPPQEPNPHSPAPCPNGS----- 505
Qy 541 QAYSGDYMEPEKCAPLLPPPPQNSVPHYAEADIVTLQGVGGNTYAVPALPPGAVGDGP 600
Db 505 -AYSGDYMEPEKCAPLLPPPPQNSVPHYAEADIVTLQGVGGNTYAVPALPPGAVGDGP 563
Qy 601 PRVDFPRSRRLFKELGEGQFGEVHLCEVDSPODLVSLDFPLNVRKGHPLLVAVKILRPD 660
Db 564 PRVDFPRSRRLFKELGEGQFGEVHLCEVDSPODLVSLDFPLNVRKGHPLLVAVKILRPD 623
Qy 661 ATKARNDFLEVKIMSLKDPNRIIRLLGVCVQDDPLCMITDMYENGDLNQLSAHQLED 720
Db 624 ATKARNDFLEVKIMSLKDPNRIIRLLGVCVQDDPLCMITDMYENGDLNQLSAHQLED 693
Qy 721 KAAGCAPGDGGAAGPTISYPMMLHVAQAQASGMRYLATLNFVHRDLATNCLVGENFTI 780
Db 694 KAAGCAPGDGGAAGPTISYPMMLHVAQAQASGMRYLATLNFVHRDLATNCLVGENFTI 743
Qy 781 KIADFGMSRNLVYAGDYRVQGRAVLPIRWMAWECILMGKFTTASDYWAFGVTLWEVLMCL 840
Db 744 KIADFGMSRNLVYAGDYRVQGRAVLPIRWMAWECILMGKFTTASDYWAFGVTLWEVLMCL 803
Qy 841 RAQPFQQLTDEQVIENAGEFPRDQGRQVYLSRPPACQGLYELMLRCWSRESQRPFSQ 900
Db 804 RAQPFQQLTDEQVIENAGEFPRDQGRQVYLSRPPACQGLYELMLRCWSRESQRPFSQ 863
Qy 901 LHRFLAEDALNTV 913
Db 864 LHRFLAEDALNTV 876
RESULT 6
W34673
ID W34673 standard; Protein; 882 AA.
AC W34673:
DT 17-FEB-1998 (first entry)
DE Human mammary carcinoma kinase 10 (MCK-10) splice variant 1.
KW Mammary carcinoma kinase; MCK-10; receptor tyrosine kinase;
KW proliferative disease; cancer; insulin receptor family;
KW tyrosine kinase neurotrophin receptor; MCK-10 activity;
KW neurological disorder; aberrant expression.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Peptide 1..18
FT /label= signal_sequence
FT Protein 19..919
FT /note= "mature_protein"
FT Domain 31..185
FT /label= Discoidin_I_like_domain
FT Cleavage_site 304..307
FT /label= endopeptidase_furin
FT /note= "putative precursor cleavage site"
FT Region 48..439
FT /label= transmembrane_region
FT Binding_site 580..590
FT /label= ATP_binding_motif
FT Modified_site 765..766
FT /label= autophosphorylation_sites
FT /note= "putative"
FT Modified_site 761
FT /label= autophosphorylation_site
FT /note= "putative"
FT Binding_site 807..810
FT /label= binding_motif_for_P13_kinase
FT /note= "binding motif for phosphatidylinositol 3'
FT Binding_site 795..795
FT /label= potential_substrate_binding_site
FT Region 26..42
FT /note= "antibody recognition sequence NTalpha"
FT Region 309..321
FT /note= "antibody recognition sequence NTbeta"

FT Region 865..882
/note= "antibody recognition sequence Crtbeta"
PN US5677144-A.
PD 14-OCT-1997.
PF 08-NOV-1994; 336343
PR 16-NOV-1993; US-153397.
PA (ALVE/) ALVES F H E.
PI (ULR/) ULLRICH A.
PI ALVES FHE, ULLRICH A;
DR WPI: 97-511869/47.
PT Truncated receptor tyrosine kinase CCK-2 - and nucleic acid coding
PT for it, useful for cancer diagnosis
PS Disclosure: Page -: 70pp; English.
CC The present sequence represents a splice variant of a mammary
CC carcinoma kinase (MCK-10). This kinase belongs to a novel family
CC of receptor tyrosine kinases, and expression is associated with
CC proliferative diseases such as cancer. The MCK-10 receptor tyrosine
CC kinase has extensive sequence similarity to the insulin receptor family.
CC The MCK-10 gene was isolated by PCR using 2 degenerate oligonucleotide
CC primer pools, using a template cDNA synthesised by reverse transcription
CC of poly-A RNA from the human mammary carcinoma cell line MCF7. The
CC amplified PCR product was used to screen human foetal brain and placental
CC libraries, from which the present splice variant was isolated. This
CC splice variant does not possess amino acids 505-541 of MCK-10 (W34672).
CC The sequence represented by amino acids 548-558 may be important, as
CC deletion of this motif in the activin receptor serine/threonine kinase
CC results in reduced ligand binding affinity. MCK-10 is expressed in brain
CC tissue, and the protein shares homology with the tyrosine kinase
CC neurotrophin receptor. Modulation of MCK-10 activity therefore may be used
CC for treatment of neurological disorders. MCK-10 is also expressed in a
CC variety of cancer cell lines and tumour tissue. The nucleotide sequence
CC of MCK-10, or parts of it, can be used for diagnostic purposes to detect
CC aberrant expression of MCK-10 genes. Inhibitors of MCK-10 (or splice
CC variants) receptor activity may have therapeutic value in the treatment
CC of diseases such as cancer.
CC note: the present sequence does not appear in the specification, but was
CC created using information provided.
SQ Sequence 882 AA;

Query Match 94.9%; Score 4651.5; DB 1; Length 882;
Best Local Similarity 95.2%; Pred. No. 0;
Matches 875; Conservative 0; Mismatches 1; Indels 43; Gaps 2;
QY 1 MGPEALSSLLLLLVASGDADMKGHFDPKARYALGMQDRTPDSDISASSWSNSTAAR 60
Db 1 MGPEALSSLLLLLVASGDADMKGHFDPKARYALGMQDRTPDSDISASSWSNSTAAR 60
QY 61 HSRLESSDGDGAWCPAGSVFPKKEEYLQVDLQRLHLVALVGTQGRHAGGLKEFSRSLR 120
Db 61 HSRLESSDGDGAWCPAGSVFPKKEEYLQVDLQRLHLVALVGTQGRHAGGLKEFSRSLR 120
QY 121 RYSRQGRRWGKRWGQGVISGNEPDGEGVVLKDLGPPWVARLVRFYPRADRVMSVCLRV 180
Db 121 RYSRQGRRWGKRWGQGVISGNEPDGEGVVLKDLGPPWVARLVRFYPRADRVMSVCLRV 180
QY 181 ELYGCLWRDGLLSYTAPEGQTMVLSYAVYNDSTYDHTVGLQYGGGLQADGVVGLDD 240
Db 181 ELYGCLWRDGLLSYTAPEGQTMVLSYAVYNDSTYDHTVGLQYGGGLQADGVVGLDD 240
QY 241 FRKSOELRVNPGYDYVGNHSHFSSGYVMEFEFDLRAFAQMVHCHNMHTLGLARLPGG 300
Db 241 FRKSOELRVNPGYDYVGNHSHFSSGYVMEFEFDLRAFAQMVHCHNMHTLGLARLPGG 300
QY 301 VECFRFRGPAWAGEPEMRHNLGNLGNLGPRAVAVPLGGRVAREFLQCRFLFAGPWLIFS 360
Db 301 VECFRFRGPAWAGEPEMRHNLGNLGNLGPRAVAVPLGGRVAREFLQCRFLFAGPWLIFS 360
QY 361 EISFISDVYNNSSPALGGTFPPAPWPPGPPPTNFSSLELEPRGQOPVAKPGGSAITLI 420
Db 361 EISFISDVYNNSSPALGGTFPPAPWPPGPPPTNFSSLELEPRGQOPVAKPGGSAITLI 420
QY 421 GCIVAILLLLLLIIALLMLRLHWRLLSKARRVLEELITVHLSVPGDITILINRPGPRE 480


```

Db 421 GCLVAILLLLLIIALMLWRLHWRLLSKAERVLSEELTVHLSVPGDTILINNRPGRPE 480
QY 481 PPYQEPFRPRGNPPHSAFCVPGNSALLSNPARYLLLATYAPRPGPGPTPAWAKPTNT 540
Db 481 PPYQEPFRPRGNPPHSAFCVPGNS----- 505
QY 541 QAYSGDYMEPEKPGAPLLPPPPQNSVPHYAEADIVTLQGTGNTYAVPALPGVAGDGP 600
Db 505 -AYSGDYMEPEKPGAPLLPPPPQNSVPHYAEADIVTLQGTGNTYAVPALPGVAGDGP 563
QY 601 PRYDFRSRLRFEKLGEGQFGEVHLCEVDSPODLVSLDFPLNVRKGHPLLVAVKILRPD 660
Db 564 PRYDFRSRLRFEKLGEGQFGEVHLCEVDSPODLVSLDFPLNVRKGHPLLVAVKILRPD 623
QY 661 ATKNA-----RNDLFKEYKIMSRLKDPNIIRLLGVCVODDPLCMITDYMENGDLNQFLS 714
Db 624 ATKNASFSLFRNDFLKEYKIMSRLKDPNIIRLLGVCVODDPLCMITDYMENGDLNQFLS 583
QY 715 AHOLEDKAAGAPGDCQAAGQPTISYPMLLHVAQAQASGMRYLATLNFVHRDLATRNCLV 774
Db 684 AHOLEDKAAGAPGDCQAAGQPTISYPMLLHVAQAQASGMRYLATLNFVHRDLATRNCLV 743
QY 775 GENFTIKIADFGMSRLYAGDYRVQGRAVLPIRMMAWECILMGKFTTASDVWAFGVTLW 834
Db 744 GENFTIKIADFGMSRLYAGDYRVQGRAVLPIRMMAWECILMGKFTTASDVWAFGVTLW 803
QY 835 EVLMLCRAQFPQGLTDEQVIENAGFEFRDQGRQVLYSRPPACPGQGLYELMLRCWSESEQ 894
Db 804 EVLMLCRAQFPQGLTDEQVIENAGFEFRDQGRQVLYSRPPACPGQGLYELMLRCWSESEQ 863
QY 895 RPPFSQHLRFLAEDALNTV 913
Db 864 RPPFSQHLRFLAEDALNTV 882

RESULT 7
W34674
ID W34674 standard; Protein: 563 AA.
AC W34674;
DT 17-FEB-1998 (first entry)
DE Human mammary carcinoma kinase 10 (MCK-10) splice variant 2.
KW Mammary carcinoma kinase; MCK-10; receptor tyrosine kinase;
KW proliferative disease; cancer; insulin receptor family;
KW tyrosine kinase neurotrophin receptor; MCK-10 activity;
KW neurological disorder; aberrant expression.
OS Homo sapiens.
FH Key Location/Qualifiers
FT 1..18
FT Peptide /label= signal_sequence
FT Protein 19..919
FT /note= "mature_protein"
FT Domain 31..185
FT /label= Discoidin_I_like_domain
FT Cleavage_site 304..307
FT /label= endopeptidase_furin
FT /note= "putative precursor cleavage site"
FT Region 48..439
FT Binding_site 617..627
FT /label= ATP_binding_motif
FT Modified_site 797..798
FT /label= autophosphorylation_sites
FT /note= "putative"
FT Modified_site 793
FT /label= autophosphorylation_site
FT /note= "putative"
FT Binding_site 839..842
FT /label= binding_motif_for_pl3_kinase
FT /note= "binding motif for phosphatidylinositol 3' kinase"
FT Binding_site 827..827
FT /label= potential_substrate_binding_site

```

```

FT Binding_site 506..509
FT /label= putative_receptor_binding_site_for_SHC
FT /note= "SHC is an oncogenic SH2 domain containing molecule"
FT Binding_site 510..513
FT /label= GTPase_activity_protein_binding_site
FT /note= "putative"
FT Region 26..42
FT /note= "antibody recognition sequence NTalpha"
FT Region 309..321
FT /note= "antibody recognition sequence NTbeta"
FT Region 897..913
FT /note= "antibody recognition sequence CTbeta"
PN US5677144-A.
PD 14-OCT-1997.
PF 08-NOV-1994; 336343.
PR 16-NOV-1993; US-153397.
PA (ALVE)/ ALVES F H E.
PA (ULLR)/ ULLRICH A.
PI Alves FHE, Ullrich A;
DR WPI: 97-511869/47.
PT Truncated receptor tyrosine kinase CCK-2 - and nucleic acid coding
PT for it, useful for cancer diagnosis
PT Disclosure; Page -: 70pp; English.
CC The present sequence represents a splice variant of a mammary
CC carcinoma kinase (MCK-10). This kinase belongs to a novel family
CC of receptor tyrosine kinases, and expression is associated with
CC proliferative diseases such as cancer. The MCK-10 receptor tyrosine
CC kinase has extensive sequence similarity to the insulin receptor family.
CC The MCK-10 gene was isolated by PCR using 2 degenerate oligonucleotide
CC primer pools, using a template cDNA synthesised by reverse transcription
CC of poly-A RNA from the human mammary carcinoma cell line MCF7. The
CC amplified PCR product was used to screen human foetal brain and placental
CC libraries, from which the present splice variant was isolated. This
CC splice variant does not possess amino acids 666-671 of MCK-10 (W34672).
CC The sequence represented by amino acids 585-595 may be important, as
CC deletion of this motif in the activin receptor serine/threonine kinase
CC results in reduced ligand binding affinity. MCK-10 is expressed in brain
CC tissue, and the protein shares homology with the tyrosine kinase
CC neurotrophin receptor. Modulation of MCK-10 activity therefore may be used
CC for treatment of neurological disorders. MCK-10 is also expressed in a
CC variety of cancer cell lines and tumour tissue. The nucleotide sequence
CC of MCK-10, or parts of it, can be used for diagnostic purposes to detect
CC aberrant expression of MCK-10 genes. Inhibitors of MCK-10 (or splice
CC variants) receptor activity may have therapeutic value in the treatment
CC of diseases such as cancer.
CC note: the present sequence does not appear in the specification, but was
CC created using information provided.
SQ Sequence : 563 AA;

```

Query Match 61.4%; Score 3011; DB 1; Length 563;
 Best Local Similarity 99.8%; Pred. No. 2e-231;
 Matches 562; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 351 LFAGPMLLFSEISFISDVNNSSPALGGTFPPAPWPPGPPPTNFSSLEPRGQGPVAK 410
Db 1 LFAGPMLLFSEISFISDVNNSSPALGGTFPPAPWPPGPPPTNFSSLEPRGQGPVAK 60
QY 411 PEGSPTAILIGCLVAIIILLLLLIIALMLWRLHWRLLSKAERVLSEELTVHLSVPGDTI 470
Db 61 AEGSPTAILIGCLVAIIILLLLLIIALMLWRLHWRLLSKAERVLSEELTVHLSVPGDTI 120
QY 471 LINNRPGPREPPYQEPFRPRGNPPHSAFCVPGNSALLSNPARYLLLATYAPRPGPGPP 530
Db 121 LINNRPGPREPPYQEPFRPRGNPPHSAFCVPGNSALLSNPARYLLLATYAPRPGPGPP 180
QY 531 TPWAKPTNTQAYSGDYMEPEKPGAPLLPPPPQNSVPHYAEADIVTLQGTGNTYAVPA 590
Db 181 TPWAKPTNTQAYSGDYMEPEKPGAPLLPPPPQNSVPHYAEADIVTLQGTGNTYAVPA 240
QY 591 LPPGAVGDGPVRVDFRSLRFEKLGEGQFGEVHLCEVDSPODLVSLDFPLNVRKGHPL 650

```


PT derived vectors, transformed cells, proteins and antibodies, useful
PT for diagnosis and treatment of proliferative and nervous system
PS diseases and for screening modulators
PS Disclosure: Page 74-77; 115pp; English.
CC A member of the mammary carcinoma kinase 10 (MCK-10) receptor
CC tyrosine kinase family was identified using a PCR (with two
CC degenerate oligo primer pools based on conserved sequences of the
CC kinase domains of receptor tyrosine kinases) and cDNA prep. from
CC colonic adenocarcinoma RNA. The nt sequence of the novel receptor,
CC designated CCK-2, is given in Q92523 and the deduced AA sequence in
CC R75505. Analysis of CCK-2 nt and AA sequences indicated significant
CC homology with MCK-10 throughout the extracellular, transmembrane
CC and intracellular regions. The regions of homology extend into the
CC N-terminus consensus sequence to the discoidin I like family of
CC proteins. CCK-2 was predominantly found in all stromal cells
CC whereas MCK-10 expression was strongly confined to neoplastic
CC cells themselves. Between the two RTKs, the juxtamembrane region
CC is the region of most extensive sequence divergence.
SQ Sequence 855 AA;

Query Match 49.3%; Score 2415; DB 1; Length 855;
Best Local Similarity 52.1%; Pred No. 8.8e-184;
Matches 492; Conservative 118; Mismatches 227; Indels 98; Gaps 15;

QY 3- PEALSLLLLLVAGSDADMKGHDFPAKCRVYALGMQDRTIPDSISASSWSWSTAARHS 62
DB 5 PRMLLVLELLPTLS---SAKAQVNPACRYPLGMSGGOIPDEDITASSOWSESTAAYG 61
QY 63 RLESSGDGAWCPAGSVFPKE-EYQLVDLQRLHLVALVGTGCRHAGGLGKEFFSRYLRL 121
DB 62 RLDSEGDGAWCPAIEPVPDDLKEFLQIDLHTLHTLTLTGCRHAGGHGIEFAPMYKIN 121
QY 122 YSDGRWRMGKDRWGQEVISNEDGEGVYLDLGPVAVRLVRYPRADRVMSCLARVE 181
DB 122 YSDGRWRISWRNKGKQVLDGNSNYDIFLKDLEPIVAVRFVPTVDHSMNVMCRVE 181
QY 182 LYGLWRDGLSYTAPVGTMYL--SEAVYLDNDYDHTVGLQYGLGLQADGVVGLD 239
DB 182 LYGCWLDGLSVYNAPAGQOQVLPGGSIYLDNDYVDG-AVGYSMTGLQTLTGVSGLD 240
QY 240 DFRKSOELRWPGDYVYVGNHSIFSSGYVEMEFEDRLRAFAMQVHCNMHTLGLARLP 299
DB 241 DFTQTHEYHWMPGYDYVGNRESATNGYIEIMFEDRIENFTTMKYHCNMFPAKGVKIFK 300
QY 300 GVECFRRGPMAWECEPMRHNGLGDPRAVSVPLGVRAREFLOCRFLFAGPMLLF 359
DB 301 EVOCYF-RSEASEWEPNAISFFLVLDVDPVPSARFVTPVPLHHRWSAICKQYHFAWTMMF 359
QY 360 SEISFISD-VYNNSSPALGGTTPPAPWPPGPPPTNFSSLELEPRGQOPVAKPEGSTAI 418
DB 360 SEITFOSDAAMYNNSEAL-----PTSP-----MAPTYDPLAKVDSNTRI 400
QY 419 LICGLVAIILLILLIITALLMLRWLRLLSKARRVLEELIVHLSVPGDITLINR--P 476
DB 401 LICGLVAIIFILLIAIIVILWQFQWOKLEKASRRMLDDEMTVSLSPSSSMFNRRSS 460
QY 477 GPREP-----PPYQEPFRGNPHPSAPCPVNGSALLISNPAYRLLLATYARP 523
DB 461 SPSEQSNSTYDIFRLRDYQEP-----SRLRLKLPF----- 495
QY 524 PRGPGPTTFAWAKPTNTQAYSGDMEPEKPGAPLLPPPPQNSVPHYAEADIVTLQGVGG 583
DB 495 -----APGEESGSCGVKVPQVSGP-----EGVPHYAEADIVNLOGVGG 535
QY 584 NTAVPALPGVAGDGPVRY-DPPRRLRFEKELGSGQFGEVHLCEVDSQDVLSDPL 642
DB 536 NTYSVPAVMTDLLSGKDVAVEEPRLKLTAFKEKLGSGQFGEVHLCEVEGMEKFKDKDFAL 595
QY 643 NVKRGHPLLVAKILRPDATTNARNDFLKEVKTMSRLKDPNIRLLGVCVQDDPLCMITD 702
DB 596 DVSAQPVLVAVKRLRADANKARNDFLKEIKIMSLKDPNIIHLSVLCITDDPLCMITE 655

QY 703 YMENGDLNQFLSAHQLEDKAAEGAPGDGQAAOQTTSYPMLLHVAQAQIASGMRYLATLNF 762
DB 656 YMENGDLNQFLSRHE-----PPNSSSSDVRTVSYTNLKPMATQIASGMKYLSSLNF 706
QY 763 VHRDLATRNCLVGENFTIKIADFGMSRNLVAGDIYRVOGRAVLPIRMMAWECILMGKFTT 822
DB 707 VHRDLATRNCLVGNKTYIKIADFGMSRNLVSGDYIYIQGRAVLPIRMMSWESILLGKFTT 766
QY 823 ASDVWAFGVTLWEVLMCLCRAOPFGQLTDEQVIENAGFEFFRDQGRQVYLSRPPACPOGLYE 882
DB 767 ASDVWAFGVTLWETFFTCQEPYSQLSDEQVIENAGFEFFRDQGRQVYLSRPPACPOSVYK 826
QY 883 LMLRCWSESEQRPPFSQLHRFLAE 907
DB 827 LMLSCWRRTKRNPSQOEIHLULLQ 851
RESULT 10
W34671 ID W34671 standard; Protein; 855 AA.
AC W34671;
DT 16-FEB-1998 (first entry)
DE CCK-2, a human mammary carcinoma kinase 10 MCK-10) family member.
KW Mammary carcinoma kinase; MCK-10; CCK-2; receptor tyrosine kinase;
KW Proliferative disease; cancer; MCK-10 activity; aberrant expression.
OS Homo sapiens.
PH Key Location/Qualifiers
FT Modified_site 121 /note= "N-glycosylated"
FT Modified_site 213 /note= "N-glycosylated"
FT Modified_site 261 /note= "N-glycosylated"
FT Modified_site 280 /note= "N-glycosylated"
FT Modified_site 328 /note= "N-glycosylated"
FT Modified_site 372 /note= "N-glycosylated"
FT Modified_site 503 /note= "N-glycosylated"
FT Modified_site /note= "putative autophosphorlation and substrate binding site"
FT Modified_site 736 /note= "putative autophosphorlation and substrate binding site"
FT Modified_site 740 /note= "putative autophosphorlation and substrate binding site"
FT Modified_site 741 /note= "putative autophosphorlation and substrate binding site"
FT Modified_site 813 /note= "putative autophosphorlation and substrate binding site"
FT Modified_site 825 /note= "putative autophosphorlation and substrate binding site"
FT Region 400..421 /label= transmembrane_region
FT Domain 30..185 /note= "putative"
FT Binding_site 433..438 /label= Discoidin_I_like_domain
FT /label= protein_kinase_C_binding_site
FT /note= "putative"
US5677144-A.
14-OCT-1997.
PD 08-NOV-1994; 336343.
PR 16-NOV-1993; US-153397.
PA (ALVE/) ALVES F H E.
PA (ULLR/) ULLRICH A.
PI Alves FHE, Ullrich A;
WPI; 97-511869/47.

DR N-PSDB; T93784.
PT truncated receptor tyrosine kinase CCK-2 - and nucleic acid coding
PS for it, useful for cancer diagnosis
PS Claim 5; Fig 3; 70pp; English.
CC The present sequence represents the amino acid sequence of human CCK-2, a
CC member of the mammary carcinoma kinase 10 (MCK-10, W34572) family of
CC receptor tyrosine kinases. The protein contains a remarkably high
CC number of proline residues arranged as PXXP or PXXP repeats, suggesting
CC a random coil structure for the hydrophilic juxtamembrane region. This
CC region is probably a major domain for interactions with cellular
CC substrates and other regulatory proteins. Expression of CCK-2 is
CC associated with proliferative diseases such as cancer. The CCK-2 gene was
CC identified by PCR and a cDNA prepared from colonic adenocarcinoma RNA.
CC CCK-2 is expressed in a wide variety of cancer cell lines and tumour
CC tissue. The CCK-2 nucleic acids can be used for diagnostic purposes to
CC detect aberrant expression of CCK-2 genes. Engineered cell lines,
CC containing recombinant vectors with the present sequence, are useful for
CC producing infectious retroviral particles. The cell lines may also be
CC used to evaluate and screen drugs involved in CCK-2 activation and
CC regulation. 855 AA;
SQ Sequence

Query Match 49.3%; Score 2415; DB 1; Length 855;
Best Local Similarity 52.1%; Pred. No. 8.8e-184;
Matches 482; Conservative 118; Mismatches 227; Indels 98; Gaps 15;

Qy 3 PEALSSLLLLLVASGDADMGHFDPAKRYALCMQDRTIPDSDISASSWSSTAAHRS 62
Db 5 PRMLLVLLPILS---SAKAQVNPACRYPLGMSGGQIPDEDITASSQWSESTAAKYG 61

Qy 63 RLESSDGDGAGCAGSVPFKE-EYLVQVLDLQRLHLVALVGTQGRHAGLGKFSRSYRLR 121
Db 62 RLDSEGDGAGCPEIPVDPDLKEFLQIDLHTLFTLVGTQGRHAGHGIEFAPWKIN 121

Qy 122 YSRDGRRMWGHKDRWGQEVISGNEDEPGVVLKDLGPPMVARLVRFYPRADRVMSVCLRVE 181
Db 122 YSRDGRTRWISWRNRHGKQVLDGNSNPYDIFLKDLEPPVAFVRFPIVTDHSMVCMRVE 181

Qy 182 LYGLWDRGLLSYAPVGTQMYL--SEAVYLDNSTYDGHVTGGLGOLADGVVGLD 239
Db 182 LYGCWLDGLVSYNAPAGQVFLPGSSIIYLDNSVDYD-AVGSMTGGLQGLDGVSGLD 240

Qy 240 DFRKSQELRVWPGDYVGVWSNHSFSSGYVMEFEFRLRAFAQMVHCHNNMHTLGARLP 299
Db 241 DFTQTHEYVWPGDYVGVWRNESATNGYIEIMEFDRIRFTTMKVHCNNMFAKGVKIF 300

Qy 300 GVECRFRGPAMWEGEPMRHNLGGLDPRARAVSVPLGGRVAFLOCFEAGFWLLF 359
Db 301 EVQCIF-RSEASEWEPNAISFPLVLDVNPFSARFVTVPLHHRWASAIKCOYHFDATWMMF 359

Qy 360 SEISFISD-VYNNSSPALGTFPPAPWPPGPPPTNFSLELEPRGQOPVAKPESGPTAI 418
Db 360 SEITFQSDAAMYNSEAL---PTSP-----MPTTYDPMKVDSDNTRI 400

Qy 419 LIGCLVAIILLIILIAIMLWRLHWRLLSKAERVLEELTVHLSVPGDTILINNR--P 476
Db 401 LIGCLVAIIFILLIIVILWROFQWKMLEKASRMDDWTVSLSPDSSMFNNRNS 460

Qy 477 GPREP-----PPQEPFRPNPHSAPCPVNGSALLSNPAYRLLIATYARP 523
Db 461 SPSSQGSNSTYDRITFPLRPDYQEP-----SRILKLEP----- 495

Qy 524 PRGCPPTPAWAKPTNTQAYSGDYMEPEKPGAPLLPPPPQNSVPHYAEADIVTLQGVG 583
Db 495 -----APGEESGCSGVVQVPSGP-----EGVPHYAEADIVNLQGVG 535

Qy 584 NTVAVPALPGVAGDGPVRV-DEPRSLRKEKELGEGQFGEVHLCVDSQDVLVSLDFPL 642
Db 536 NTVSVPAVTMDLLSGKDVAVEEFPRKLTTEKELGEGQFGEVHLCVDSQDVLVSLDFPL 595

Qy 643 NVKRGHPLLVAVKTLRPDANKNENDFLKEVKIMSLRKLDPNIIILLCVQVDDPCLMID 702
Db 643 NVKRGHPLLVAVKTLRPDANKNENDFLKEVKIMSLRKLDPNIIILLCVQVDDPCLMID 702

Db 596 DVSANQPVLVAVKMLRADANKNARNDFLKEIKIMSRLKDPNIIHLLSVCIITDDPLCMITE 655
Qy 703 YMENGDLNQFLSAHOLEDKAAEGAPGCGOAAAGPTISYPMLLHVAAQIASGMRYLATLNF 762
Db 656 YMENGDLNQFLSRHE-----PPNSSSDVTRVSTNLFKFWATQIASGMKYLSSLNF 706

Qy 763 VHRDLATRNCLVGENFTIKIADFGMSRNLYAGDYRYVQGRAVLPIRMWMECILMGKFTT 822
Db 707 VHRDLATRNCLVGNKYTIKIDFGMSRNLYSGDYRYIQGRAVLPIRMWMSWESILLGKFTT 766

Qy 823 ASDWAFAGVTLWEVLMLCRAQPPFQGLTDEQVIENAGEFFRDGGRQVILSRPPACQGLYE 882
Db 767 ASDWAFAGVTLWEVTFQEQEYISQLSDEQVIENAGEFFRDGGRQVILSRPPACQGLYE 826

Qy 883 LMLRCWSRESEORPPFSOLHRLAE 907
Db 827 LMLSCWRDRDTRNRPFSQIHLILLQ 851

RESULT 11
W77114
ID W77114 standard; Protein; 855 AA.
AC W77114;
DT 16-NOV-1998 (first entry)
DE Discoidin domain receptor 2 protein.
KW Discoidin domain receptor; transformatin; metastasis; collagen; ss;
KW Cleidocranial dysplasia; Sickler syndrome; extracellular matrix; MMP-1.
OS Homo sapiens.
PN W09834954-A2.
PD 13-AUG-1998.
PF 05-FEB-1998; CA0093.
PR 06-FEB-1997; US-041578.
PA (MOUN) MOUNT SINAI HOSPITAL CORP.
PI Pawson A. Vogel W;
DR WPI: 98-447188/38.
DR N-PSDB; V48292.
PT Novel ligands of discoidin domain receptor tyrosine kinase,
PT especially collagen - useful for treating e.g. metastasis,
PT cleidocranial dysplasia or Sickler syndrome
PS Disclosure: Fig 22a; 115pp; English.
CC The discoidin domain receptor (DDR) can be used to identify and evaluate
CC substances which affect DDR receptor tyrosine kinase signalling pathways
CC in the cell. Compounds which modulate such signalling pathways can be
CC used to alter transformation or metastasis in mammals, to treat
CC conditions involving structural or functional deregulation of collagens,
CC e.g. Cleidocranial dysplasia or Sickler syndrome, conditions requiring
CC modulation of extracellular matrix synthesis, degradation or remodelling,
CC or to treat conditions needing modulation of MMP-1 expression such as
CC wound healing.
SQ Sequence 855 AA;

Query Match 49.3%; Score 2415; DB 1; Length 855;
Best Local Similarity 52.1%; Pred. No. 8.8e-184;
Matches 482; Conservative 118; Mismatches 227; Indels 98; Gaps 15;

Qy 3 PEALSSLLLLLVASGDADMGHFDPAKRYALCMQDRTIPDSDISASSWSSTAAHRS 62
Db 5 PRMLLVLLPILS---SAKAQVNPACRYPLGMSGGQIPDEDITASSQWSESTAAKYG 61

Qy 63 RLESSDGDGAGCAGSVPFKE-EYLVQVLDLQRLHLVALVGTQGRHAGLGKFSRSYRLR 121
Db 62 RLDSEGDGAGCPEIPVDPDLKEFLQIDLHTLFTLVGTQGRHAGHGIEFAPWKIN 121

Qy 122 YSRDGRRMWGHKDRWGQEVISGNEDEPGVVLKDLGPPMVARLVRFYPRADRVMSVCLRVE 181
Db 122 YSRDGRTRWISWRNRHGKQVLDGNSNPYDIFLKDLEPPVAFVRFPIVTDHSMVCMRVE 181

Qy 182 LYGLWDRGLLSYAPVGTQMYL--SEAVYLDNSTYDGHVTGGLGOLADGVVGLD 239
Db 182 LYGCWLDGLVSYNAPAGQVFLPGSSIIYLDNSVDYD-AVGSMTGGLQGLDGVSGLD 240

Qy 240 DFRKSQELRVWPGDYVGVWSNHSFSSGYVMEFEFRLRAFAQMVHCHNNMHTLGARLP 299
Db 241 DFTQTHEYVWPGDYVGVWRNESATNGYIEIMEFDRIRFTTMKVHCNNMFAKGVKIF 300

Qy 300 GVECRFRGPAMWEGEPMRHNLGGLDPRARAVSVPLGGRVAFLOCFEAGFWLLF 359
Db 301 EVQCIF-RSEASEWEPNAISFPLVLDVNPFSARFVTVPLHHRWASAIKCOYHFDATWMMF 359

Qy 360 SEISFISD-VYNNSSPALGTFPPAPWPPGPPPTNFSLELEPRGQOPVAKPESGPTAI 418
Db 360 SEITFQSDAAMYNSEAL---PTSP-----MPTTYDPMKVDSDNTRI 400

Qy 419 LIGCLVAIILLIILIAIMLWRLHWRLLSKAERVLEELTVHLSVPGDTILINNR--P 476
Db 401 LIGCLVAIIFILLIIVILWROFQWKMLEKASRMDDWTVSLSPDSSMFNNRNS 460

Qy 477 GPREP-----PPQEPFRPNPHSAPCPVNGSALLSNPAYRLLIATYARP 523
Db 461 SPSSQGSNSTYDRITFPLRPDYQEP-----SRILKLEP----- 495

Qy 524 PRGCPPTPAWAKPTNTQAYSGDYMEPEKPGAPLLPPPPQNSVPHYAEADIVTLQGVG 583
Db 495 -----APGEESGCSGVVQVPSGP-----EGVPHYAEADIVNLQGVG 535

Qy 584 NTVAVPALPGVAGDGPVRV-DEPRSLRKEKELGEGQFGEVHLCVDSQDVLVSLDFPL 642
Db 536 NTVSVPAVTMDLLSGKDVAVEEFPRKLTTEKELGEGQFGEVHLCVDSQDVLVSLDFPL 595

Qy 643 NVKRGHPLLVAVKTLRPDANKNENDFLKEVKIMSLRKLDPNIIILLCVQVDDPCLMID 702
Db 643 NVKRGHPLLVAVKTLRPDANKNENDFLKEVKIMSLRKLDPNIIILLCVQVDDPCLMID 702

Db 241 DTQTHEYVWPXYDVGWRNESATNGIETMFEFDRINFTTKMKVHCNNFAGVKRIFK 300
 QY 300 GVECFRRGPMAMWEGEPWRHNLGNLGDPRARAVSVPLGGRVARFLQCRFLFAGPWLLE 359
 Db 301 EVQCYF-RSEASEWEPNAISFPLVLDVNPASRFVPLHHRMASAIKCOQHFADTWMP 359
 QY 360 SEISFISD-VYNNSSPALGGTFPPAPWPPGPPPTNFSSSLEPRGGOOPVAKPEGSPAI 418
 Db 360 SEITFQSDAANYNSEAL-----PTSP-----MAPTVDPMKLKVDSDNTRI 400
 QY 419 LGICLVAILLLILLIALLMLARLHWRLLSKAERVLLEELTVHLSVPGDTILINNR--P 476
 Db 401 LGICLVAILFILLALIVILLWRQFQWKLKASRMULDDEMTVSLSPSSSMFNNRNS 460
 QY 477 GPREP-----PPYQEPGRPNPSPHAPCPVNGSALLSNPVARLLLATYARP 523
 Db 461 SPSEQGSNSTYDRIFPLRPDQEP-----SRLIRKLPEF----- 495
 QY 524 PRGPGPTPAWAKPTNTQAYSGDYMEPEKPCAPLLPPPPQNSVPHYAEADIVTLQGV 583
 Db 495 -----APGEEESGCGVVKVPQPSGP-----EGVPHYAEADIVNLQGV 535
 QY 584 NTYAVPAPPGAVGDPGRV-DEPRSLRREKELGEGQFGEVHLCVDSQDLVSLDFPL 642
 Db 536 NTYSPVATVMDLLSGKDVAVEEFPKLLTFKELGEGQFGEVHLCVDSQDLVSLDFPL 595
 QY 643 NVKRGHPLLVAVKILRPDATKNARNDFLKEVKIMSRLLKDPNIIIRLLGVCVDDPLCMITD 702
 Db 596 DVSANQPVAVKMLRADANKARNDFLKEVKIMSRLLKDPNIIIRLLGVCVDDPLCMITD 655
 QY 703 YWENGDLNQFLSAHQLEDKAAEGAPGQQAAGQPTISYPMLLHVAAQIASGRYLAFLNF 762
 Db 656 YWENGDLNQFLSRHE-----PPNSSSDVTVSYTNLKFATQIASGMKYLSSNF 706
 QY 763 VHRDLATNCLVGENFTIKIADFGMSRNLVAGDYRYQGRVAVLPIRMWAWECILLMGKFTT 822
 Db 707 VHRDLATNCLVGNFTIKIADFGMSRNLVSGDYRYQGRVAVLPIRMWAWESILLGKFTT 766
 QY 823 ASDVWAFVGLTWELMLCRAQFQGLTDEQVIENAGFEFFRQGRVYLSRPPACQGLYE 882
 Db 767 ASDVWAFVGLTWETFTFCQEQPYSQLSDEQVIENAGFEFFRQGRVYLSRPPACQGLYE 826
 QY 883 LMLRCWSESEQRPPFSLHRLFAE 907
 Db 827 LMLSCWRDRTKNRPSFQBIHLLLIQ 851

RESULT 12

W79152
 ID W79152 standard; Protein; 854 AA.
 AC W79152;
 DT 19-NOV-1998 (first entry)
 DE Receptor protein tyrosine kinase (PTK) subtype tyro-10.
 KW PTK; receptor; protein tyrosine kinase; brain tissue.
 OS Rattus sp.
 PN US811516-A.
 PD 22-SEP-1998.
 PF 15-MAY-1992; 456647.
 PR 02-JUN-1995; US-884486.
 PR 02-MAY-1994; US-237401.
 PR 02-JUN-1995; US-456647.
 PA (SALK) SALK INST BIOLOGICAL STUDIES.
 PI Lai CHC, Lemke GE.
 DR WPI; 98-53039/45.
 DR N-PSDB; V55895.
 PT Receptor protein tyrosine kinase polypeptide, tyro-3 - preferably expressed in brain tissue
 PS Example 2; Columns 53-60; 46pp; English.
 CC This represents a novel receptor protein tyrosine kinase (PTK) polypeptide subtype tyro-10. The invention provides polynucleotide sequences encoding novel PTK polypeptide subtypes tyro-1 to tyro-13.
 CC The PTK subtypes are found expressed predominantly in the brain tissue.

SQ Sequence 854 AA;

Query Match 49.2%; Score 2413; DB 1; Length 854;
 Best Local Similarity 52.3%; Pred. No. 1.3e-183;
 Matches 481; Conservative 119; Mismatches 220; Indels 100; Gaps 15;

QY 9 LLLLLLVAGSDADMKGHFDPAKRYALGMDRTIPDSDISASSSSSDSTAARHSLESD 68
 Db 10 VLLLLLLILGSA--KAQVNPACRPLGMSGCHIDEDITASSQNSSESTAARIGLDEE 67
 QY 69 GGAWCPAGSVFPKE--EYLOVDLQRLHLVALVGTQGRHAGGLGKFEFSRYRLRSRGR 127
 Db 68 GGAWCPETPVQDDLEKFLQDLRLTFLITLVGTQGRHAGGHGTEFAPMYKINSRDGS 127
 QY 128 RWMGKDKRWGOEIVSNGEDPEGVWLKDLGPPMVARLVREYPRADRVMSVCLRVLYGLW 187
 Db 128 RWMGKDKRWGOEIVSNGEDPEGVWLKDLGPPMVARLVREYPRADRVMSVCLRVLYGLW 187
 QY 188 RDGLLSYTPAVGQTMYL--SEAVYLNDSYDGTHTVGGLOGLQGLADGVGLDDFRKSQ 245
 Db 188 LDGLVSYNAPAGQQFVLPQGSIIYLNDSYDGTHTVGGLOGLQGLADGVGLDDFRKSQ 246
 QY 246 ELRWPGDYVYVWNSHSSFGVYEMEFEDRLRAFQAMQVHCNNMHTLGLARLPGGVECRF 305
 Db 247 EYVWPGDYVYVWNSHSSFGVYEMEFEDRLRAFQAMQVHCNNMHTLGLARLPGGVECRF 306
 QY 306 RGPAMAWEGEPWRHNLGNLGDPRARAVSVPLGGRVARFLQCRFLFAGPWLLESEISFI 365
 Db 307 -RSEASEWEPNAISFPLVLDVNPASRFVPLHHRMASAIKCOQHFADTWMPSEITFQ 365
 QY 366 SD--VYNNSSPALGGTFPPAPWPPGPPPTNFSSSLEPRGGOOPVAKPEGSPAILIGCL 423
 Db 366 SDAAAYNNS-----GALPTSP-----MAPTVDPMKLKVDSDNTRIIGCL 405
 QY 424 VAILLLLLIIIALMLWRHLRRLLSKAERVLLEELTVHLSVPGDTIILNRR-----PGPR 479
 Db 406 VAIIFILLAIIVILWRQFQWKLKASRMULDDEMTVSLSPSSSMFNNRNSSPSEQ 465
 QY 480 EP-----PYQEPGRPNPSPHAPCPVNGSALLSNPVARLLLATYARPGRPG 528
 Db 466 ESNSTYDRIFPLRPDQEP-----SRLIRKLPEF----- 495
 QY 529 PPTPAWAKPTNTQAYSGDYMEPEKPCAPLLPPPPQNSVPHYAEADIVTLQGVGTGNTYAV 588
 Db 495 -----APGEEESGCGVVKVPQPSGP-----EGVPHYAEADIVNLQGVGTGNTYAV 540
 QY 589 PALPPGAVGDPGRV-DFPRSLRREKELGEGQFGEVHLCVDSQDLVSLDFPLNVRKG 647
 Db 541 PAVTMDLLSGKDVAVEEFPKLLTFKELGEGQFGEVHLCVDSQDLVSLDFPLNVRKG 600
 QY 648 HPLLVAVKILRPDATKNARNDFLKEVKIMSRLLKDPNIIIRLLGVCVDDPLCMITDYMENG 707
 Db 601 QPVLVAVKMLRADANKARNDFLKEVKIMSRLLKDPNIIIRLLGVCVDDPLCMITDYMENG 660
 QY 708 DLNQFLSAHQLEDKAAEGAPGQQAAGQPTISYPMLLHVAAQIASGRYLAFLNFVHRDL 767
 Db 661 DLNQFLSRHEPLSSCSDA-----TVSYANLKFATQIASGMKYLSSNLFVHRDL 710
 QY 768 ATRNCLVGENFTIKIADFGMSRNLVAGDYRYQGRVAVLPIRMWAWECILMGKFTTASDVW 827
 Db 711 ATRNCLVGNKNTIKIADFGMSRNLVSGDYRYQGRVAVLPIRMWAWESILLKFTTASDVW 770
 QY 828 AFGVTLWELMLCRAQFQGLTDEQVIENAGFEFFRQGRVYLSRPPACQGLYELMLRC 887
 Db 771 AFGVTLWETFTFCQEQPYSQLSDEQVIENAGFEFFRQGRVYLSRPPACQGLYELMLRC 830
 QY 888 WSESEQRPPFSLHRLFAE 907
 Db 831 WRRETKHRPSFQBIHLLLIQ 850

RESULT 13

W81409

ID W81409 standard; Protein; 854 AA.
 AC W81409.
 DT 22-JAN-1999 (first entry)
 DE Receptor protein tyrosine kinase (PTK) subtype tyro-10.
 KW PTK; receptor; protein tyrosine kinase; recombinant; grafting;
 KW diagnosis; tumour; skin transplant; connective tissue; tyro-10.
 OS Rattus sp.
 PN US5937448-A.
 PD 17-NOV-1998.
 PF 02-MAY-1994; 237401.
 PR 15-MAY-1992; US-884486.
 PR 02-MAY-1994; US-237401.
 PA (SALK) SALK INST BIOLOGICAL STUDIES.
 PI Lai CHC. Lemke GE.
 DR WPI: 99-023436/02.
 DR N-PSDB; V65317.
 PT Nucleic acids encoding protein tyrosine kinase subtypes - for
 PT identification of new subtypes and treatment of diseases associated
 PT with the kinase
 PS Claim 10: Columns 53-58: 47pp: English.
 CC This represents a receptor protein tyrosine kinase (PTK) subtype tyro-10.
 CC The invention provides sequences V65308 to V65313, V65315, and V65317 to
 CC V65319 that encode proteins having a tyrosine kinase domain and a tissue
 CC expression pattern of a receptor PTK subtype selected from tyro-1,
 CC tyro-2, tyro-3, tyro-4, tyro-5, tyro-6, tyro-8, tyro-10, tyro-11, and
 CC tyro-12, respectively. The polynucleotides are useful for the detection
 CC of tyrosine kinase domain sequences and detection of tissue expression
 CC patterns of PTK subtypes. The cDNAs can also be injected into oocytes,
 CC the protein expressed, and expression products screened for using
 CC antibodies against tyrosine kinase epitopes. These subtypes sequences can
 CC be used for the design of oligonucleotides, for use in amplification
 CC reactions to isolate other subtype sequences. These detection protocols
 CC are used in the diagnosis of diseases associated with (receptor) PTKs.
 CC Recombinant vectors expressing the subtypes can be used to treat related
 CC diseases e.g. tumours, by introduction of the vectors into skin
 CC transplants, then grafting these into the connective tissue of the
 CC dermis, thus specifically targeting tumours as the proteins are released
 CC from the matrix.
 CC Sequence 854 AA;
 SQ

Query Match 49.2%; Score 2413; DB 1; Length 854;

Best Local Similarity 52.3%; Pred. No. 1.3e-183;

Matches 481; Conservative 119; Mismatches 220; Indels 100; Gaps 15;

QY 9 LLLLLLVASGDADMKGHFPAKCRYALGMQDRTIPSDISASSWSDSSTAARHRSLESD 68
 DB 10 VLLLLLLILGSA--KAQVNPACIRYPLGMSGHGPEDITASSQWSESTAAYGRLDSEE 67
 QY 69 GDGAWCPAGSVFPKE-EYELQVDLQRLHLVALVGTQGRHAGGLGKFFRSYRLRYSRDGR 127
 DB 68 GDGAWCEIPVQPDLEKEFLQIDLRLFTLTLVGTQGRHAGHGIEFAPWKYKINSRDSG 127
 QY 128 RWGKWDKRWQGVISGNDPEGVVLKDLGPPMVARLVRYPPRADRVMSVCLRVLYGCLW 187
 DB 128 RWISWRNRHGKQVLDGNSNEYDFELKDEPIVAREVRLPIVTHSMVMCMRVLYGCVW 187
 QY 188 RDGLLSYTPVGTQMYL--SEAYLVNDSTYDGHVTGLQVGLGQLADGVVGLDDFRKSQ 245
 DB 188 LDGLVSNAPAGQOYVLPGGGSIYLLNDSDYDG-AVGYSMTEGQUTDGVGSLDDFTQTH 246
 QY 246 ELRWPGYDYVGNHNSFFSGYVEMEFEDRLRAFAMQVHCNMMHTLGLARLPGGVCEFR 305
 DB 247 EYHWVPCYDVGVRNESATNGFIEFEDRIIRNFTTMKVHCNMMFAKGVKIFKEVQCYF 306
 QY 306 RRGPAWAGEPMHNLGNLGDPRARAVSVPLGGVRVAREFLQCRFLFAGPWLLEFSISFI 365
 DB 307 -RSEASEWETAVYFPLVDVNPASRFVTVPLHHRMASAIAKCOYHFADTWMFSEITFQ 365
 QY 366 SD--VVNNSPALGGTTPPAPWPPGPPPTNFSLSLEPRGQOPVAKPESPTAILIGCL 423
 DB 366 SDAAMYNNS-----GALTPSP-----MAPTYDPMULKVDSDNTRILIGCL 405

QY 424 VAITLLLLIIALMLRLHWRLLSKAERVRLEELTVHLSVPGDTILNNR-----PGPR 479
 DB 406 VAIIFILLAIIVILWRQFWQKMLEKASRMDDDEMTVLSLPSESMFNNRNSPSEQ 465
 QY 480 EP-----PPYQEPFRPNPPHSAFCVPGVNGSALLSNPAYLLATYARPPRGG 528
 DB 466 ESNSTYDRIFPLRPDYQEP-----SRLIRKLPEF----- 495
 QY 529 PPTPAWAKPTNTQAYSCDYMEPEKPGAPLPPPPQNSVPHYAADIIVTLOGVTGGNTYAV 588
 DB 495 -----APGEEESGCGVVKPAQNGP-----EGVPHYAADIIVNLOGVTGGNTYCV 540
 QY 589 PALPPGAVGDPGRV-DFPRSRIRFKELGEGQFGEVHLCEVDSPQDLVSLDPLNVRKG 647
 DB 541 PAYTMDLLSGKDVAVEEFPRKLLAFKELGEGQFGEVHLCEVEGMEKFKDKDFALDSAN 600
 QY 648 HPLVAVKILRPDATKNARNDFLKEVKIMSRKDPNIIRLLGVCVQDDPLCMITDMENG 707
 DB 601 QPVLVAVKMLRADANKNARNDFLKEIKIMSRKDPNIIRLLAVCIITDPLCMITEYMEG 660
 QY 708 DLNQFLSAHQLEDKAAEGAPGDCGAAGPTISYPMLLHVAQAQIASGMRYLATINEVHRDL 767
 DB 661 DLNQFLSRHREPLSCSSDA-----TVSYANLKFEMATQASGMKYLSSNFVHRDL 710
 QY 768 ATRNCLVGENFTIKIADFGMSRNLVAGDYRVQGRAVLPPIRMMAWECILMGKFTASDVW 827
 DB 711 ATRNCLVGNKVTIKIADFGMSRNLVAGDYRVQGRAVLPPIRMMSWESILLGKFTASDVW 770
 QY 828 AFGVTLMVLMCRAQPGQITDEQVITENAGFEFRGQGVYLSRPPACQGYELMLRC 887
 DB 771 AFGVTLMETTFCEQYPSQLSDQVIENTGEFFRDGROIYLPQALCPDSPYKMLSC 830
 QY 888 WRESEQRPPESQLHRELA 907
 DB 831 WRRETKHRPSFOEIHLLLLQ 850
 RESULT 14
 R54089
 ID R54089 standard; Protein; 650 AA.
 AC R54089;
 DT 03-FEB-1995 (first entry)
 DE Partial sequence of tyrosine kinase receptor protein.
 KW Tyrosine kinase; receptor; proto-oncogene; trk; detection;
 KW diagnosis; antibody; treatment; tumour; antisense.
 OS Homo sapiens.
 FH Key. Location/Qualifiers
 FT modified_site 8
 FT modified_site 56 /note= "Potential N-linked glycosylation site."
 FT modified_site 75 /note= "Potential N-linked glycosylation site."
 FT modified_site 167 /note= "Potential N-linked glycosylation site."
 FT modified_site 195..216 /note= "Potential N-linked glycosylation site."
 FT domain 365..370 /label= Transmembrane domain.
 FT region /label= ATP binding region.
 FT /note= "Highly conserved among protein tyrosine
 FT kinase enzymes."
 PN DE4239817-A.
 PD 01-JUN-1994.
 PF 26-NOV-1992; 239817.
 PR 26-NOV-1992; DE-239817.
 PA (CHEM-) CHEMOTHERAPEUTISCHES FORSCHUNG.
 PI Holtrich U, Ruebsamen-waigmann H, Strebhardt K;
 DR WPI: 94-184380/23.
 DR N-PSDB; 064158.
 PT New protein tyrosine kinase and related nucleic acid - vectors,
 PT transformed cells, etc., useful for diagnosis and treatment of
 PT tumours

PS Claim 1; Page 8; 9pp; German.
CC The gene encoding this polypeptide is related to the trk
CC proto-oncogene. Antibodies against the polypeptide are useful for
CC diagnosis and for the treatment of tumours. The antibodies may also
CC be radiolabelled or coupled to a cytotoxin for destruction of cancer
CC cells. Antisense nucleic acid can be used to inhibit gene
CC expression. 650 AA;
SQ Sequence 650 AA;

Query Match 37.2%; Score 1825; DB 1; Length 650;
Best Local Similarity 51.5%; Pred. No. 4.9e-137;
Matches 369; Conservative 89; Mismatches 167; Indels 92; Gaps 12;

QY 208 VYLDSTYDGHVTGGYQYGLGQGLADGVVGLDDFRKSELVWPGYDYVWGNHFSFSSGY 267
Db 5 IYLDNSYVDG-AVGYSMTEGLGQLTDGVSGLDDFTQTHEYHVWPGYDYVWGNHFSFSSGY 63
QY 268 VEMFEFEDRLRAFQAOVHCNHNHTLGARLPGGVCEFRFRGPMAMWEGEPMRHNLGGNLG 327
Db 64 IEMFEFDRIRNFTTKVHCNHNFAKGVKIFKEVOCYF-RSEASEWEPNAISFPPLVLDV 122
QY 328 DPRARAVSVPLGGRVAFQCRFLFAGPMLLFSEIFSD-VVNNSSPALGGTFPPAPWW 386
Db 123 NPSARFVTVPLHRMASAIKQHFADTWMFSEITFQSDAAMYNSEAL----PTSP-- 177
QY 387 PGPPPTNFSLEPRGQOPVAKPEGSPTAILIGLCLVAIILLLLIIALMLWLHWRRL 446
Db 177 -----MAPTTYDPMKLKVDSDNTRILIGLCLVAIIFILLAIIVILWRQFQWM 223
QY 447 LSAERVLLEETVHLVSPGDTILNRR--PGPREP-----PYQEPFRPG 491
Db 224 LEKASRLMDDEMTVSLSPDSMFNNRSPSQSGSNSTYDRIFPLRPDYQEP---- 280
QY 492 NPHSAPCVNGSALLSNPARYLLATYARPPGPTPAWAKPTNTQAYSGDYMEPE 551
Db 280 -----SRLKLPDEF-----APGEEBSCSGVYKPV 305
QY 552 KPGAPLPPPPQNSPHYAEADITVLQGVGTNTYAVPALPGAVGDPGRPY-DPPRSRL 610
Db 306 QPSGP-----EGVPHYAEADIVNLQGVGTNTYVPAVTWDLGSKDVAVEEPRKLL 358
QY 611 RFKEKLEGQFGEVHLCGEVSDSPDLVSLQFPLNVRKGHPDLLVAVKILRPDATKNARNDPL 670
Db 359 TFKEKLEGQFGEVHLCGEVMEKFKDKDFALDSANQPVLVAVKMLRADANKNARNDPL 418
QY 671 KEVKINSLKDPNIIRLLGVCVDDPLCMITDYMENGDLNQFLSAHQLEDKAAEGAPGDG 730
Db 419 KEIKINSLKADPNIIHLLSVCIITDPLCMITETMENGDLNQFLSRHE-----PPNS 469
QY 731 QAAQGGPTISYPMLLHVAAQIASGMRYLATLNFVHRDLATRNCLVGENFTIKIADFGMSRN 790
Db 470 SSSDVRTVSYTNLKWATQIASCMKYLSSILNFVHRDLATRNCLVGENFTIKIADFGMSRN 529
QY 791 LYAGDYRVQGRAVLPIRMWAWECILMGKFTTASDVWAGVTLWEVLMICRAQPGQLTFD 850
Db 530 LYSGDYRIQGRAVLPIRMWAWESILLGKFTTASDVWAGVTLWEVTLWETFTFCQEPYSLSD 589
QY 851 EQVTENAGEFFRDGROVLYSRPPACPGIYELMLRCWRESFQRPFPQSLHFLAE 907
Db 590 EQVIENTGEFFRDGROVLYSRPPACPGIYELMLRCWRESFQRPFPQSLHFLAE 907

RESULT 15
R81630
ID R81630 standard; Protein; 822 AA.
AC R81630;
DT 31-Mar-1996 (first entry)
DE Human trkB receptor protein.
KW trkB receptor; tyrosine-kinase; enzyme; protease; inflammation; pain;
KW diagnosis; neurotrophic factor; kidney; lung; psychiatric disorder.
OS Homo sapiens.
FH Location/Qualifiers

FT misc_difference 67.70 /note="potential N-linked glycosylation site"
FT misc_difference 95.98 /note="potential N-linked glycosylation site"
FT misc_difference 121.124 /note="potential N-linked glycosylation site"
FT misc_difference 178.181 /note="potential N-linked glycosylation site"
FT misc_difference 205.208 /note="potential N-linked glycosylation site"
FT misc_difference 241.244 /note="potential N-linked glycosylation site"
FT misc_difference 254.257 /note="potential N-linked glycosylation site"
FT misc_difference 280.283 /note="potential N-linked glycosylation site"
FT misc_difference 325.328 /note="potential N-linked glycosylation site"
FT misc_difference 338.341 /note="potential N-linked glycosylation site"
FT misc_difference 412.415 /note="potential N-linked glycosylation site"
FT domain 431.454 /note="transmembrane domain"
FT domain 544.807 /note="tyrosine-kinase domain"
FT misc_difference 466 /note="splice site for truncated trkC"
FN W09525795-A1.
PD 28-SEP-1995.
PF 17-MAR-1995; U03426.
PR 18-MAR-1994; US-215139.
PR 05-AUG-1994; US-286846.
PR 20-DEC-1994; US-359705.
PR (GETH) GENENTECH INC.
PA Presta LG, Shelton DL, Urfer R;
DR WPI; 95-344616/44.
DR N-PSDB; T00689.
DR New human trkB and trkC poly:peptide(s) and fusion proteins contg.
DR them - also DNA, vectors and transformed cells useful in treatment
PT and diagnosis of abnormal neurotrophic factor expression, e.g.
PT inflammatory pain
PS Claim 8; Fig.1A-B; 117pp; English.
CC This sequence may be expressed recombinantly for the production
CC of human trkC receptor, and to detect or amplify trkC genes. The
CC encoded protein may be used as a reagent in kinase receptor
CC activation assays, and therapeutically in diseases associated with
CC over or under expression of neurotrophic factor (e.g. pain of
CC inflammation, kidney, lung, cardiovascular or psychiatric disorders
CC and some sorts of tumours). At the indicated splice site, AA 436-
CC 791 may be replaced by the sequence FVLHFKIPLDG' (R81624).
SQ Sequence 822 AA;

Query Match 13.4%; Score 658; DB 1; Length 822;
Best Local Similarity 47.4%; Pred. No. 4.2e-44;
Matches 144; Conservative 45; Mismatches 79; Indels 36; Gaps 9;

QY 607 RSLRPFKEKLEGQFGEVHLCGEVSDSPDLVSLQFPLNVRKGHPDLLVAVKILRPDATK 663
Db 535 RHNVLRRELGEAGAGVFLAECYNLCPEQ-----KILVAVKTLK-DASD 579

QY 664 NARNDLKEVKIMSLKDPNIIRLLGVCVDDPLCMITDYMENGDLNQFLSAHQLED-KA 722
Db 580 NARKDFRHEALLNLQHEHIVKEYGVCVGDPLIMVFEYMKHGDGLNFKLFRANGPDVLM 639
QY 723 AEGAPGDGQAQGGPT-ISYPMLLHVAAQIASGMRYLATLNFVHRDLATRNCLVGENFTIK 781
Db 640 AEGNP-----FTELTQSOHLHIAQIAGVYVLAQSHFVHRDLATRNCLVGENLLVK 691
QY 782 IADFGMSRNLYAGDYRVQGRAVLPIRMWAWECILMGKFTTASDVWAGVTLWEVLMCLR 841
Db 692 IGDFGMSRDVSTDYRVGGHTMLPIRMWPPESIMYKFTTESDWSLSGVVLEIFTYK 751

Qy 842 AOPFGQLTDEQVIENAGEFFRDQGRQVYLSRPPACPOGLYELMLRCWSRESEQRPFSQL 901
Db 752 -QPWYQLSNNEVIECI-----TQGR--VLQRPTCPQEVYELMLGCVQREPHMRKNIKGI 803
Qy 902 HRFL 905
Db 804 HTLL 807

Search completed: November 3, 1999, 22:52:09
Job time: 33331 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run On: November 4, 1999, 07:05:18 ; Search time 752.95 Seconds
(without alignments)
3135.829 Million cell updates/sec

Title: US-08-170-558-7

Perfect score: 1197

Sequence: 1 GATGCTGACATGAGGGACA.....CCGAGGGGAGCCCGACCGCC 1197

Scoring table: IDENTITY_NUC

Searched: 2546578 seqs, 986266752 residues

Database :

EST:*

1: em_est1:*

2: em_est2:*

3: em_est3:*

4: em_est4:*

5: em_est5:*

6: em_est6:*

7: em_est7:*

8: em_est8:*

9: em_est9:*

10: em_est10:*

11: em_est11:*

12: em_est12:*

13: em_est13:*

14: em_est14:*

15: em_est15:*

16: em_est16:*

17: em_est17:*

18: em_est18:*

19: em_est19:*

20: gb_est1:*

21: gb_est2:*

22: gb_est3:*

23: gb_est4:*

24: gb_est5:*

25: gb_est6:*

26: gb_est7:*

27: gb_est8:*

28: gb_est9:*

29: gb_est10:*

30: gb_est11:*

31: gb_est12:*

32: gb_est13:*

33: gb_est14:*

34: gb_est15:*

35: gb_est16:*

36: gb_est17:*

37: gb_est18:*

38: gb_est19:*

39: gb_est20:*

40: gb_est21:*

41: gb_est22:*

42: gb_est23:*

43: gb_est24:*

44: gb_est25:*

45: gb_est26:*

46: gb_est27:*

47: gb_est28:*

48: gb_est29:*

49: gb_est30:*

50: gb_est31:*

51: gb_est32:*

52: em_est20:*

53: em_est21:*

54: em_est22:*

55: em_est23:*

56: em_est24:*

57: em_est25:*

58: em_est26:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	474.4	39.6	772	54	HSN008101	Al043251 Homo sapi
2	314.8	26.3	433	28	AA073943	AA073943 mm97hl2.r
C 3	210.8	17.6	568	36	AA622035	AA622035 nm23b08.s
4	177.8	14.9	257	54	HSN011191	Al046341 Homo sapi
C 5	119.4	10.0	245	44	AI300297	AI300297 gm69hl2.x
6	116.8	9.8	235	23	R85358	R85358 yo39hl2.r1
7	101.6	8.5	316	20	T28305	T28305 EST37127.Hu
8	73.2	6.1	82	23	H38476	H38476 YP6C07.r1
9	46.2	3.9	495	37	AA734307	AA734307 vt26f03.r
10	45.8	3.8	515	49	AV000273	AV000273 AV000273
11	44.6	3.7	520	30	AA221816	AA221816 my28f05.r
12	44.6	3.7	503	35	AA597143	AA597143 vo35f03.r
13	44.6	3.7	484	37	AA691209	AA691209 vt34e06.r
14	44.6	3.7	507	37	AA691386	AA691386 vut5a10.r
15	44.6	3.7	518	37	AA692234	AA692234 vt23h01.r
16	44.6	3.7	596	39	AA815999	AA815999 vr14b12.r
17	44.2	3.7	335	38	AA781615	AA781615 ai56f01.s
C 18	44.2	3.7	530	39	AA857320	AA857320 oh97a11.s
C 19	44.2	3.7	441	40	AA991406	AA991406 os52d12.s
20	44.2	3.7	243	49	AU058693	AU058693 AU058693
21	44.2	3.7	243	49	AU058845	AU058845 AU058845
22	44.2	3.7	243	49	AU058851	AU058851 AU058851
C 23	44	3.7	501	36	C77167	C77167 C77167 Mous
24	43.8	3.7	459	23	D52589	D52589 HUM082810B
25	43.8	3.7	243	49	AU059967	AU059967 AU059967
C 26	43.2	3.6	453	44	AI289917	AI289917 qv02el2.x
C 27	43	3.6	427	28	AA084238	AA084238 zn04h06.s
C 28	43	3.6	565	28	AA108143	AA108143 ml96b10.r
C 29	43	3.6	489	34	AA499317	AA499317 vi92h05.r
C 30	43	3.6	500	37	AA734118	AA734118 vv22f10.r
C 31	43	3.6	736	39	AA822411	AA822411 vw37d09.r
C 32	43	3.6	781	42	AI092944	AI092944 qa81c01.x
C 33	43	3.6	722	42	AI098196	AI098196 ue31c12.x
C 34	43	3.6	495	47	AI538451	AI538451 td06d12.x
C 35	43	3.6	456	49	AV001579	AV001579 AV001579
C 36	42.8	3.6	466	29	AA150564	AA150564 zo89a04.s
C 37	42.6	3.6	419	34	AA468397	AA468397 nc78g07.s
C 38	42.6	3.6	568	35	AA587115	AA587115 nn70a10.s
C 39	42.6	3.6	510	38	AA738127	AA738127 nx16e01.s
C 40	42.6	3.6	447	39	AA838639	AA838639 ac40e11.s
C 41	42.6	3.6	492	39	AA857507	AA857507 oh98e07.s
C 42	42.6	3.6	439	41	AI002243	AI002243 or72d12.s
C 43	42.6	3.5	377	23	D52591	D52591 HUM082C03B
44	42.2	3.5	342	23	D54105	D54105 HUM128H08B
45	42.2	3.5				

ALIGNMENTS

RESULT 1
HSN008101
ID HSN008101 standard; RNA; EST; 772 BP.
XX
AC AL043251;
XX
SV AL043251.1
XX

NI e1401623

XX 12-MAR-1999 (Rel. 59, Created)

DT 12-MAR-1999 (Rel. 59, Last updated, Version 1)

XX

DE Homo sapiens mRNA; EST DKFZp434K0223_r1 (from clone

DE DKFZp434K0223)

XX

KW EST: expressed sequence tag.

XX

OS Homo sapiens (human)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria;

OC Primates; Catarrhini; Hominoidea; Homo.

XX

RN [1]

RP 1-772

RA Blum H., Bauersachs S., Mewes W., Gassenhuber J., Wiemann S.;

RT Submitted (12-MAR-1999) to the EMBL/GenBank/DBJ databases.

RL MIPS, Am Klopferspitz 18a D-82152 Martinsried, GERMANY

XX

CC Clone from S. Wiemann, sequenced by LMU within the cDNA

CC sequencing consortium of the German Genome Project

CC No s1 sequence available

CC This clone is available at the RZPD in Berlin

CC Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059

CC Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de

XX

FH Key Location/Qualifiers

FT source

FT 1..772

FT /db_xref="taxon:9606"

FT /organism="Homo sapiens"

FT /clone="DKFZp434K0223"

FT /clone_lib="434 (synonym: htes3). Vector pSPORT1; host

FT DH10B; sites NotI + SalI"

FT /dev_stage="adult"

FT /tissue_type="testis"

XX

XX Sequence 772 BP; 118 A; 250 C; 260 G; 143 T; 1 other;

XX

Query Match 39.6%; Score 474.4; DB 54; Length 772;

Best Local Similarity 99.2%; Pred. No. 2.6e-98;

Matches 497; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 1 GATGCTGACATGAGGACATTTTGTCTGCAAGTCCCGTATGCGCTGGGATGCAG 60

DB 256 GATGCTGACATGAGGACATTTTGTCTGCAAGTCCCGTATGCGCTGGGATGCAG 315

QY 61 GACCGGACCATCCAGACAGTGCATCTCTGCTTCAGCTCTGCTCAGATTCACACTGCC 120

DB 316 GACCGGACCATCCAGACAGTGCATCTCTGCTTCAGCTCTGCTCAGATTCACACTGCC 375

QY 121 GCCCGCCACAGCAGGTTGGAGACAGTGCAGGGGATGCGGCTGTTGCCCGCAGGGTCG 180

DB 376 GCCCGCCACAGCAGGTTGGAGACAGTGCAGGGGATGCGGCTGTTGCCCGCAGGGTCG 435

QY 181 GTGTTTCCCAAGGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 240

DB 436 GTGTTTCCCAAGGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 495

QY 241 CTGGTGGGACCCAGGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 300

DB 496 CTGGTGGGACCCAGGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 555

QY 301 CGGCTGCTTACTCCCGGAGTGTGCGCGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 360

DB 556 CGGCTGCTTACTCCCGGAGTGTGCGCGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 615

QY 361 GAGGTGATCTCAGGCAATGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 420

DB 616 GAGGTGATCTCAGGCAATGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 674

QY 421 ATGTTTGGCCGACTGGTCTGCTTACCCCGGGGTGACCGGGGTATGAGGCTGTGCTG 480

DB 675 ATGTTTGGCCGACTGGTCTGCTTACCCCGGGGTGACCGGGGTATGAGGCTGTGCTG 734

QY 481 CGGGTAGAGCTCTATGGCTGC 501

DB 735 CGGGTAGA-CTCTATGGCTGC 754

RESULT 2

AA073943 433 bp mRNA EST 15-FEB-1997

LOCUS mm97h12.r1 Stratagene mouse heart (#937316) Mus musculus cDNA clone

DEFINITION IMAGE:536423 5' similar to TR:G1161063 G1161063 RECEPTOR KINASE. ;

RNA sequence.

ACCESSION AA073943

NID 91595690

VERSION AA073943.1 GI:1595690

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 433)

AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.

TITLE The WashU-HMI Mouse EST Project

JOURNAL Unpublished (1996)

COMMENT On Sep 12, 1996 this sequence version replaced gi:1400901.

Contact: Marra M/Mouse EST Project

WashU-HMI Mouse EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@wustl.edu

This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:323359

Possible reversed clone: similarity on wrong strand

Seq primer: -28ml3 rev1 ET from AmerSham

High quality sequence stop: 429.

Location/Qualifiers

1..433

/organism="Mus musculus"

/strain="NIH/Swiss"

/db_xref="taxon:10090"

/clone="IMAGE:536423"

/clone_lib="Stratagene mouse heart (#937316)"

/sex="pooled"

/tissue_type="heart"

/dev_stage="13 day embryos"

/lab_host="SOLR (kanamycin resistant)"

/note="Organ: heart; Vector: pBluescript SK-; Site:1: EcoRI; Site:2: XhoI; Cloned unidirectionally. Primer: Oligo dt. 93 pooled NIH/Swiss 13 day embryo hearts. Average insert size: 1.0 kb; Uni-ZAP XR Vector: -5' adaptor sequence: 5' GAATTCGGCAGCAG 3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTT 3"

BASE COUNT 78 a 103 c 155 g 96 t . 1 others

ORIGIN

Query Match 26.3%; Score 314.8; DB 28; Length 433;

Best Local Similarity 84.7%; Pred. No. 3.7e-62;

Matches 365; Conservative 0; Mismatches 63; Indels 3; Gaps 1;

QY 294 GAGCTACCGGCTGGTACTTACCCGGGATGTCGCCGCTGGTGGCTGGAAGGACCGCTG 353

```

Db 1 GAACTATCGTTCGGTACTCCGAGATGCCCGCTGGATGACTGGAAGACCGCTG 60
QY 354 GGGTCAGGAGTGATCTCAGGCAATGAGGACCTGAGGAGGAGTGTCTCAAGGACCTTGG 413
Db 61 GGGACAGGAGTGATTTGGGTACAGAGATCCCGGGGAGTAGTCTGAAGACCTTGG 120
QY 414 GCGCCCATGTTGCGGACGATGTTTCGCTTCTACCCCGGGGTGACCGGGTCAATGAGCT 473
Db 121 ACCCGCATGTTGCGGCGGTGCTGCTTCTACCCAGGGCTGACCGGATCATGAGT 180
QY 474 CTGCTGCGGATAGCTCTATGCTGCTCTGAGGAGGATGACCTCTCTTACACCGC 533
Db 181 CTGCTTTCGGGTGGAGCTCTATGCTGCTCTGCGGAGTGGAGCTCTCTCATATACAGC 240
QY 534 CCTGTGGGCGACACATGATTTATCTGAGG---CCGTGTACTCAACGACTCCACCTA 590
Db 241 CCCCGTCGAGNAGCATGACGATTTCTGAGGTGATGTTACATCTCAATGATCCACTA 300
QY 591 TGACGACATACCGTGGCGGAGCTGAGTATGAGGTGATGAGTGGCGAGCTGGCAGATGGT 650
Db 301 CGATGATATACCTGCTGAGGAGGCTGAGTATGAGTGGCGAGCTGGCGAGCTGGCAGTGGCGT 360
QY 651 GGTGGGCTGGATGACTTTAGGAGATCAGGAGCTGCGGGTCTGGCCAGGCTATGACTA 710
Db 361 GGTGGGCTGGATGATTTAGGAGATCAGGAGCTGCGGGTCTGGCCAGGCTATGACTA 420
QY 711 TGTGGGATGGA 721
Db 421 TGTGGGATGGA 431

RESULT 3
AA622035/c
LOCUS
DEFINITION
ACCESSION
NID
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

```

Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert_Strausberg@nih.gov
 Tissue procurement: Ilan Kirsch, M.D., Michael R. Emmert-Buck,
 M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www.bio.lnl.gov/bbrp/image/image.html

Insert Length: 1405 Std Error: 0.00
 Seq primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 423.

FEATURES

source
 1..568
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="3"
 /clone="IMAGE:1144695"

/clone_lib="NCI-CGAP_Col0"
 /tissue_type="colon tumor RER"
 /lab_host="DH10B"

/note="Organ: colon; Vector: pT73D-Pac (Pharmacia) with a
 modified polylinker; 1st strand cDNA was prepared from
 RER+ colon tumor, and was then primed with a Not I -
 oligo(dT) primer. Double-stranded cDNA was ligated to Eco
 RI adaptors (Pharmacia), digested with Not I and cloned
 into the Not I and Eco RI sites of the modified pT73
 vector. Library is normalized. Library was constructed by
 Bento Soares and M. Fatima Bernaldo (N-Soares4)."
 BASE COUNT 150 a 158 c 131 g 127 t 2 others
 ORIGIN

Query Match 17.68; Score 210.8; DB 36; Length 568;
 Best Local Similarity 95.88; Pred. No. 1.7e-38;
 Matches 226; Conservative 0; Mismatches 9; Indels 1; Gaps 1;
 QY 502 CTCTGGAGGATGACTCTCTTACACCCGCTGCGGAGACAAATGTTATCT 561
 Db 568 CTCTGGAGGATGACTCTCTTACACCCGCTGCGGAGACAAATGTTATCT 510
 QY 562 GAGCCGCTGACTCAACGACTCCACCTATGAGGACATACCTGGCGGAGTGCAGTAT 621
 Db 509 GAGCCGCTGACTCAACGACTCCACCTATGAGGACATACCTGGCGGAGTGCAGTAT 450
 QY 622 GGGGCTCTGGCCAGCTGGCAGATGTTGGTGGGCTGGATGACTTTAGGAAGTGCAG 681
 Db 449 GGGGCTCTGGCCAGCTGGCAGATGTTGGTGGGCTGGATGACTTTAGGAAGTGCAG 390
 QY 682 GAGCTCGGGTCTGGCCAGCTATGACTATGTTGGATGGAGCAACACAGCTTCTC 737
 Db 389 GAGCTCGGGTCTGGCCAGCTATGACTATGTTGGATGGAGCAACACAGCTTCTC 334

RESULT 4

HSM011191
 ID HSM011191 standard; RNA; EST; 257 BP.
 XX AC AL046341;
 XX SV AL046341.1
 XX NI e1404815
 XX DT 12-MAR-1999 (Rel. 59, Created)
 DT 12-MAR-1999 (Rel. 59, Last updated, Version 1)
 XX DE Homo sapiens mRNA; EST DKFZp434J247_r1 (from clone
 DE DKFZp434J247)
 XX EST; expressed sequence tag.
 KW Homo sapiens (human)
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria;
 OC Primates; Catarrhini; Hominidae; Homo.
 XX [1]
 RP 1-257
 RA Koehler K., Beyer A., Mewes W., Gassenhuber J., Wiemann S.;
 RT Submitted (12-MAR-1999) to the EMBL/GenBank/DBJ databases.
 RL MIPS, Am Klopferspitze 18a D-82152 Martinsried, GERMANY
 RL CC Clone from S. Wiemann, sequenced by BMFZ within the cDNA
 CC sequencing consortium of the German Genome Project
 CC No s1 sequence available
 CC This clone is available at the RZPD in Berlin
 CC Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 CC Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de
 XX Key Location/Qualifiers

```

FH source
FT 1. .257
FT /db_xref="taxon:9606"
FT /organism="Homo sapiens"
FT /clone="DFP243J247"
FT /clone_lib="434 (synonym: htes3). Vector pSport1; host
FT DH10B; sites NotI + SalI"
FT /dev_stage="adult"
FT /tissue_type="testis"
FT
XX
SQ Sequence 257 BP; 40 A; 97 C; 67 G; 51 T; 2 other;

Query Match 14.9%; Score 177.8; DB 54; Length 257;
Best Local Similarity 97.8%; Pred. No. 3.9e-31;
Matches 179; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1015 TTACTCTTCAGCAATCTCTTCATCTCTGATGTGTGACAAATTCCTTCGGGCACTG 1074
DB 1 TTACTCTTCAGCAATCTCTTCATCTCTGATGTGTGACAAATTCCTTCGGGCACTG 60

QY 1075 GGAGCACCCTCCCGCCAGCCCTGTGTGGCCGCTGCGCCACCTCCACCACTTGAGC 1134
DB 61 GGAGCACCCTCCCGCCAGCCCTGTGTGGCCGCTGCGCCACCTCCACCACTTGAGC 120

QY 1135 AGCTTGGAGCTGGAGCCAGAGCCAGCAGCCCGTGGCCAGCCCGAGGGGAGCCGACC 1194
DB 121 AGCTTGGAGCTGGAGCCAGAGCCAGCAGCCCGTGGCCAGCCCGAGGGGAGCCGACC 180

QY 1195 GCC 1197
DB 181 GCC 183

RESULT 5
LOCUS AI300297/c 245 bp mRNA EST 29-JAN-1999
DEFINITION qm59h12.x1 Soares_placenta_8to9weeks_2NbHP8to9w Homo sapiens CDNA
clone IMAGE:1894055.3' similar to SW:EDD1.HUMAN Q08345 EPITHELIAL
DISCOLDIN DOMAIN RECEPTOR 1 PRECURSOR ;, mRNA sequence.
ACCESSION AI300297
NID 93959643
VERSION AI300297.1 GI:3959643
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 245)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Aug 21, 1998 this sequence version replaced.

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 921 Std Error: 0.00
Seq primer: -400p from Gibco
High quality sequence stop: 59.
Location/Qualifiers
1..245
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1894055"
/clone_lib="Soares_placenta_8to9weeks_2NbHP8to9w"
/dev_stage="two placentae; one from 8 weeks and another
from 9 weeks post conception"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: Placenta; Vector: pT73D (Pharmacia) with a

modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGAGGCGCGCGATTTTTTTTTTTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library constructed by Bento Soares and
M.Fatima Bonaldo."
BASE COUNT 58 a 48 c 61 g 78 t
ORIGIN

Query Match 10.0%; Score 119.4; DB 44; Length 245;
Best Local Similarity 95.3%; Pred. No. 7.1e-18;
Matches 123; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 ATGCTGACATGAAGGACATTTTGATCTCTGCAAGTGGCGCTATGCTGCGCATGCAGG 61
DB 189 ATGCTGACATGAAGGACATTTTGATCTCTGCAAGTGGCGCTATGCTGCGCATGCAGG 130

QY 62 ACCGGACCATCCACAGATGACATCTCTGCTTCCAGCTCCTGTGTCAGATTCACATGCCG 121
DB 129 ACCGGACCATCCACAGATGACATTTTGTCTCCAGCTCCTGTGTCAGATTCACATGCCG 70

QY 122 CCCGCCACA 130
DB 69 CTCGCCCA 61

RESULT 6
LOCUS R85358 235 bp mRNA EST 14-AUG-1995
DEFINITION yo39h12.r1 Soares adult brain N2b4HB55Y Homo sapiens CDNA clone
IMAGE:180359.5' similar to SP:S37402 S37402 TRKE PROTEIN - ;, mRNA
sequence.
ACCESSION R85358
NID 9943764
VERSION R85358.1 GI:943764
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 235)
AUTHORS Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
Trevisan, E., Waterston, R., Williamson, A., Wohlmann, P. and
Wilson, R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Willson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
Insert Size: 2589
High quality sequence starts: 1
High quality sequence stops: 1
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality
Possible reversed clone: Similarity on wrong strand
Insert Length: 2589 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 1.
Location/Qualifiers
1..235
/organism="Homo sapiens"
/db_xref="GDB:3827004"
FEATURES
source
```


AUTHORS

Hillier L., Clark N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevasakis, E., Waterston, R., Williamson, A., Wohldmann, P. and Wilson, R.

TITLE

The WashU-Werck EST Project

JOURNAL

Unpublished (1995)

COMMENT

On May 5, 1995 this sequence version replaced gi:798177.

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

High quality sequence starts: 1

High quality sequence stops: 1

Source: IMAGE Consortium, LLNL

This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Trace considered overall poor quality

Possible reversed clone: similarity on wrong strand

Insert Length: 2546 Std Error: 0.00

Seq primer: M13Rp1

High quality sequence stop: 1.

FEATURES

source

Location/Qualifiers

1..82

/organism="Homo sapiens"

/db_xref="GDB:3830900"

/db_xref="taxon:9606"

/clone="IMAGE:192884"

/clone_lib="Soares adult brain N2b4HB55Y"

/sex="Male"

/dev_stage="55-year old"

/lab_host="DH10B (ampicillin resistant)"

/note="Organ: brain; Vector: pT73D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'

TGTTACCAATCTGAAGTGGAGCGCGCGCTTTTTTTTTTTT 3']

double-stranded cDNA was size selected, ligated to Eco RI

adapters (Pharmacia), digested with Not I and cloned into

the Not I and Eco RI sites of a modified pT73 vector

(Pharmacia). Library went through one round of

normalization to a Cot = 53. Library constructed by Rento

Soares and M.Fatima Donaldo. The adult brain RNA was

provided by Dr. Donald H. Gilden. Tissue was acquired

17-18 hours after death which occurred in consequence of a

ruptured aortic aneurysm. RNA was prepared from a pool of

tissues representing the following areas of the brain:

frontal, parietal, temporal and occipital cortex from the

left and right hemispheres, subcortical white matter,

basal ganglia, thalamus, cerebellum, midbrain, pons and

medulla."

BASE COUNT

13 a 32 c 20 g 13 t 4 others

ORIGIN

Query Match

Best Local Similarity 6.1%; Score 73.2; DB 23; Length 82;

Matches 75; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1071 ACTGGGAGCAGCTTCCGCCAGCCCCCTGTGGCGGCTGCCACCTCCCACTT 1130

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1 ACTTGAGGACCTTCCGCCAGCCCCCTGTGGGNCCTGCCAACCTCCCACTT 60

QY 1131 CAGCAGCTTGGAGCTGGAGCCC 1152

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 61 CAGCAGCTTGGAGCTGGAGCCC 82

RESULT

AA734307

LOCUS

DEFINITION

AA734307 495 bp mRNA EST 07-JAN-1998
vt26f03.r1 Barstead mouse proximal colon MPLRB6 Mus musculus CDNA
clone IMAGE:1164221 5' similar to gb:J03941 Mouse ferritin heavy
chain (MOUSE);, mRNA sequence.

ACCESSION

NID

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

On May 18, 1995 this sequence version replaced gi:811120.

Contact: Marra M/Mouse EST Project

WashU-HHMI Mouse EST Project

Washington University School

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:630133

Seq primer: -28ml3 rev2 ET from Amersham

High quality sequence stop: 494.

Location/Qualifiers

1..495

/organism="Mus musculus"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="IMAGE:1164221"

/clone_lib="Barstead mouse proximal colon MPLRB6"

/dev_stage="7 day juvenile"

/lab_host="DH10B"

/note="Vector: pT73D-Pac (Pharmacia) with a modified

polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA

was primed with a Not I oligo(dT) primer [5'

TGTTACCAATCTGAAGTGGAGCGCGCGCTTTTTTTTTTTT 3']

double-stranded cDNA was ligated to Eco RI

adapters (Pharmacia), digested with Not I and cloned into the

Not I and Eco RI sites of the modified pT73 vector.

Library constructed by Bob Barstead."

BASE COUNT 141 a 113 c 118 g 123 t

ORIGIN

Query Match

Best Local Similarity 3.9%; Score 46.2; DB 37; Length 495;

Matches 117; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 626 GTCTGGCCAGCTGGCAGATGGTGGTGGGCTGGATGACTTTAGGAGAGCTCAGGAGC 685

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 74 GACTGGGAGAGCGGGCTGAATGCAATGGAGTGTGCACCTGTTGAAAAGAGTGTGAATC 133

QY 686 TGGCGGCTGGCCAGGCTATGACTGTGGATGGAGCAACACACAGCTTCTCCAGTGGCT 745

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 134 AGTCACTATGGAACTGCACAACTGGCTACTGACAAGAAATGATCCCCACTTATGTGACT 193

QY 746 ATGTGAGATGGAGTTTGGTTTGACCGGCTGAGGGCTTCCAGGCTATGACAGGTCCACT 805

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 194 TCATTGAGACGTATTATCTGAGTGAACAGGTGAATCCATTAAAGAACTGGGTGACCAGC 253

QY 806 GTACACATGACACAGCTGGGAGCGGCTGCTCCCTGGGCGGTGGAAATGTCCTT 860

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 274 TGACCAACTTACGCAAGATGGGTGCCCTGAAGCTGGCATGGCAGAAATATCTCTT 328

RESULT 11

AA221816	LOCUS	AA221816	520 bp	mrna	EST	13-FEB-1997
	DEFINITION	my28f05.r1 Barstead mouse pooled organs MPlR84 Mus musculus cDNA clone IMAGE:697185 5' similar to gb:J03941 Mouse ferritin heavy chain (MOUSE);, mRNA sequence.				
	ACCESSION	AA221816				
	NID	gi1841062				
	VERSION	AA221816.1	GI:1841062			
	KEYWORDS	EST.				
	SOURCE	house mouse.				
	ORGANISM	Mus musculus				
		Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
	REFERENCE	1 (bases 1 to 520)				
	AUTHORS	Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisell,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Willson,R. and Waterston,R.				
	TITLE	The WashU-HMI Mouse EST Project				
	JOURNAL	Unpublished (1996)				
	COMMENT	On Nov 29, 1993 this sequence version replaced gi:636346.				

Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse Est Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Putative full length read
vector to vector length is 522
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 507.

```

FEATURES
source
  1. ..520
  Location/Qualifiers
    organism="Mus musculus"
    strain="FVB/N"
    db_xref="taxon:10090"
    clone="IMAGE:697185"
    clone_lib="Barstead mouse pooled organs MPLRBA"
    sex="mixed"
    tissue_type="pooled organs"
    dev_stage="7 day"
    slat_bosc="bui09"

```

/note="Organ:	pooled; Vector:	pt7T3D-Pac	(Pharmacia) with a modified polylinker;	Site_1:	EcoRI;	Site_2:	NotI;	lst
strand cDNA was primed with a Not I - oligo(dT) primer [5								
TGTTACAGATCTGAAGTGGAGCGCCGCTTTTTTTTTTTTTTTTTTTTTTTT								
3']; double-stranded cDNA was ligated to Eco RI adaptors								
[GTGTCATTGGTAGC], digested with Not I and cloned into								
the Not I and Eco RI sites of the modified pT773 vector.								
Library constructed by Bob Barstead."								
152 a	118 C	126 g	124 t					
ORIGIN								
Query Match	3.7%	Score 44.6;	DB 30;	Length 520;				
Best Local Similarity	49.4%;	Pred. No. 0.9;						
Matches 116;	Conservative	0;	Mismatches 119;	Indels	0;	Gaps	0;	
QY	626	GTCTGGGCCAGCTGGCGAGATGGTGTTGGTGGGGCTGCATGACTTTAGGAAGAGATCAGGACC	685					
Ddb	99		158					
QY	686	TGCGGGGTCTGGCCCCAGGCTATGACTATGTGTGGGATGGAGCAACAACAGACTTCTCCAGCTGGCT	745					

Not I and Eco RI sites of the modified pT73 vector.
Library constructed by Bob Barstead. "

BASE COUNT
ORIGIN

139 a 109 c 116 g 120 t

Query Match 3.7%; Score 44.6; DB 37; Length 484;
Best Local Similarity 49.4%; Pred. No. 0.87;
Matches 116; Conservative 0; Mismatches 119; Indels 0; Gaps 0;

QY 626 GTCTGGCCAGCTGCACATGGTGTGGTGGGCTGGATGACCTTTAGGAAGAGTCAGGAGC 685
DB 65 GACTGGGAGAGCGGGCTGAATGCAATGAGGTGTCACCTTGGAAAGAGGTGTAATC 124

QY 686 TCGGGTCTGGCCAGGCTATGACTATGTGGATGGAGCAACACACAGCTTCTCCAGTGGCT 745
DB 125 AGTCACTACTGGAATGCACAACTGGCTACTGACAAAGATGATCCCACTTATGTGACT 184

QY 746 ATGTGGAGATGGAGTTTGAAGTTGACCGGCTGAGGGCCCTCCAGGCTATGAGGTCCACT 805
DB 185 TCATTGAGACGTATTATCTGAGTGACACAGGTGAAATCCATTAAAGAACTGGGTGACCAAG 244

QY 806 GTAAACAATGACACAGCTGGGAGCCCTCTCCCTGGCGGGGTGGAATGTCGCTT 860
DB 245 TGACCAACTTACCAAGATGGGTGCCCTGGAAGCTGGCATGGCAGATATCTCTT 299

RESULT 14
AA691386

LOCUS
DEFINITION

AA691386 507 bp mRNA EST 16-DEC-1997

IMAGE:1180698 5', similar to gb:X12812 Murine mRNA for macrophage ferritin heavy subunit (MOUSE);, mRNA sequence.

ACCESSION
NID

AA691386 507 bp mRNA EST 16-DEC-1997

VERSION
KEYWORDS

AA691386 507 bp mRNA EST 16-DEC-1997

SOURCE
ORGANISM

AA691386 507 bp mRNA EST 16-DEC-1997

REFERENCE
AUTHORS

Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
The WashU-HHMI Mouse EST Project
Unpublished (1996)
On Sep 12, 1996 this sequence version replaced gi:1402308.

TITLE
JOURNAL
COMMENT

Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:638546
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 423.

FEATURES
SOURCE

1..507
/organism="Mus musculus"
/strain="C3H"
/db_xref="taxon:10090"
/map="925E08; 2; 2q12.3-2q14.3"
/clone="IMAGE:1180698"
/clone_lib="Barstead mouse myotubes MPLRB5"
/cell_line="C2C12"
/lab_host="DH10B"

/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTACGAATCTGAAGTGGAGCGGCCCTTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors [AATCGGATCCTTG], digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library constructed by Bob Barstead. The C2C12 cell line (available from ATCC, catalog # CRL-1772) differentiates rapidly, forming contractile myotubes and producing characteristic muscle proteins."

BASE COUNT
ORIGIN

144 a 113 c 128 g 122 t

Query Match 3.7%; Score 44.6; DB 37; Length 507;
Best Local Similarity 49.4%; Pred. No. 0.89;
Matches 116; Conservative 0; Mismatches 119; Indels 0; Gaps 0;

QY 626 GTCTGGCCAGCTGCACATGGTGTGGTGGGCTGGATGACCTTTAGGAAGAGTCAGGAGC 685
DB 87 GACTGGGAGAGCGGGCTGAATGCAATGAGGTGTCACCTTGGAAAGAGGTGTAATC 146

QY 686 TCGGGTCTGGCCAGGCTATGACTATGTGGATGGAGCAACACACAGCTTCTCCAGTGGCT 745
DB 147 AGTCACTACTGGAATGCACAACTGGCTACTGACAAAGATGATCCCACTTATGTGACT 206

QY 746 ATGTGGAGATGGAGTTTGAATGACCGGCTGAGGGCCCTCCAGGCTATCCAGGTCCACT 805
DB 207 TCATTGAGACGTATTATCTGAGTGACAGGTGAAATCCATTAAAGAACTGGGTGACCAAG 266

QY 806 GTAAACAATGACACAGCTGGGAGCCCTCTCCCTGGCGGGGTGGAATGTCGCTT 860
DB 267 TGACCAACTTACCAAGATGGGTGCCCTGGAAGCTGGCATGGCAGATATCTCTT 321

RESULT 15
AA692234

LOCUS
DEFINITION

AA692234 518 bp mRNA EST 16-DEC-1997

IMAGE:1163953 5', similar to gb:J03941 Mouse ferritin heavy chain (MOUSE);, mRNA sequence.

ACCESSION
NID

AA692234 518 bp mRNA EST 16-DEC-1997

VERSION
KEYWORDS

AA692234 518 bp mRNA EST 16-DEC-1997

SOURCE
ORGANISM

house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 518)
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
The WashU-HHMI Mouse EST Project
Unpublished (1996)
On Sep 19, 1997 this sequence version replaced gi:1517401.

TITLE
JOURNAL
COMMENT

Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:629865
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 403.
Location/Qualifiers

GenCore version 4.5

Copyright (c) 1993 - 1998 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 4, 1999, 07:29:20 ; Search time 60.16 Seconds
(without alignments)
2070.145 Million cell updates/sec

Title: US-08-170-558-7

Perfect score: 1197

Sequence: 1 GATGCTGACATGAGGACCA.....CCGAGGGAGCCCGACGCGC 1197

Scoring table: IDENTITY_NUC

Searched: 192659 seqs, 52021692 residues

Database : Issued_Patents_NA:**

1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/5C_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/5D_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS9_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1197	100.0	3637	2	US-08-445-640-3
2	1197	100.0	1197	2	US-08-445-640-7
3	1192.2	99.6	3962	2	US-08-336-343A-1
4	327.4	27.4	3157	2	US-08-336-343A-3
5	327.4	27.4	3157	2	US-08-336-343A-5
6	321	26.8	3120	3	US-08-456-647B-19
7	321	26.8	3120	3	US-08-237-401A-19
8	46.6	3.9	6402	3	US-08-670-707A-36
9	46.6	3.9	4334	3	US-08-670-707A-38
10	39.6	3.3	461	3	US-09-032-684-18
11	37.6	3.1	7218	2	US-08-232-463-14
12	37.6	3.1	432	3	US-08-642-255-48
13	37.6	3.1	756	3	US-08-642-255-50
14	37.2	3.1	1116	3	US-08-672-569-1
15	37.2	3.1	817	3	US-08-672-569-2
16	37.2	3.1	522	3	US-08-672-569-3
17	36.6	3.1	3728	1	US-08-111-939-1
18	36.2	3.0	10596	1	US-07-884-811-15
19	36.2	3.0	10596	1	US-07-885-971-15
20	36.2	3.0	10596	1	US-08-087-783A-15
21	36.2	3.0	10596	1	US-08-194-088B-15
22	36.2	3.0	10596	4	US-08-194-087-15
23	36.2	3.0	10596	5	PCT-US93-04648-15
24	36	3.0	2830	1	US-07-882-292-1
25	36	3.0	2830	5	PCT-US93-04102-1
26	35.8	3.0	2103	4	US-08-735-041A-5
27	35.8	3.0	2103	4	US-08-735-041A-7
28	35.2	2.9	7218	2	US-08-232-463-14
29	34.8	2.9	7559	3	US-08-250-848-2
30	34.6	2.9	2483	2	US-08-464-340A-3
31	34.6	2.9	5392	4	US-08-403-852D-1
32	34.6	2.9	833	4	US-08-403-852D-3
33	34.6	2.9	2483	5	PCT-US94-08449A-3
34	33.8	2.8	1594	4	US-08-955-713-1
35	33.6	2.8	13875	4	US-08-734-344-1
36	33.2	2.8	932	1	US-08-458-912-1
37	33.2	2.8	932	2	US-08-461-179-1

38 33.2 2.8 932 3 US-08-459-254-1 Sequence 1, Appli
39 33.2 2.8 932 3 US-08-459-255-1 Sequence 1, Appli
40 33 2.8 12001 3 US-08-458-568A-11 Sequence 11, Appl
41 33 2.8 2647 5 PCT-US93-06251-77 Sequence 77, Appl
42 32.8 2.7 1466 1 US-08-056-051-3 Sequence 3, Appli
43 32.8 2.7 1610 1 US-08-056-051-5 Sequence 5, Appli
44 32.8 2.7 659 1 US-07-928-611-10 Sequence 10, Appl
45 32.8 2.7 803 1 US-07-928-611-12 Sequence 12, Appl

ALIGNMENTS

RESULT 1
US-08-445-640-3
; Sequence 3, Application US/08445640
; Patent No. 5709858
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J.
; APPLICANT: Mark, Melanie R.
; APPLICANT: Scadden, David T.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Baron, Will F.
; TITLE OF INVENTION: Protein Tyrosine Kinases
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/445,640
; FILING DATE: 22-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/170558
; FILING DATE: 20-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/157563
; FILING DATE: 23-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hasak, Janet E.
; REGISTRATION NUMBER: 28,616
; REFERENCE/DOCKET NUMBER: 854C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/952-9881
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3637 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-445-640-3

Query Match 100.0%; Score 1197; DB 2; Length 3637;
Best Local Similarity 100.0%; Pred. No. 2; 7e-287;
Matches 1197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GATGCTGACATGAAGGACATTTTGATCTGCTCAAGTGCCTATGCTCGGCGATGCG 60
|||||
Db 136 GATGCTGACATGAAGGACATTTTGATCTGCTCAAGTGCCTATGCTCGGCGATGCG 195
|||||
QY 61 GACCGGACCATCCAGACAGATGACATCTCTGCTTCCAGCTCTGCTGATGATCCACTGCC 120

Db 196 GACGGACCATCCACAGACAGTACATCTCTGCTTCCAGCTCCTGGTCAAGATTCACATGCC 255
 Qy 121 CCCGCCACACAGAGTTGGAGAGCAGTACAGGGGATGGGCTGGTCCCGCCGACAGGCTCG 180
 Db 256 CCCGCCACACAGAGTTGGAGAGCAGTACAGGGGATGGGCTGGTCCCGCCGACAGGCTCG 315
 Qy 181 GTGTTTCCAGAGGAGGAGTACTTTCAGGTGATCTACACAGCTGACACTGGTGGCT 240
 Db 316 GTGTTTCCAGAGGAGGAGTACTTTCAGGTGATCTACACAGCTGACACTGGTGGCT 375
 Qy 241 CTGTGGGACCCAGGAGGAGCAGTCCGGGGCTGGGCAAGGATCTCCCGAGAGTAC 300
 Db 376 CTGTGGGACCCAGGAGGAGCAGTCCGGGGCTGGGCAAGGATCTCCCGAGAGTAC 435
 Qy 301 CGGTGCTTACTCCCGGATGGTGGCGCTGGATGGGCTGGAGAGCAGCTGGGGTCAG 360
 Db 436 CGGTGCTTACTCCCGGATGGTGGCGCTGGATGGGCTGGAGAGCAGCTGGGGTCAG 495
 Qy 361 GAGGTGATCTCAGGAATGAGAGCCTGAGGAGTGGTGTGAAGAGACTTGGGCCCGCC 420
 Db 496 GAGGTGATCTCAGGAATGAGAGCCTGAGGAGTGGTGTGAAGAGACTTGGGCCCGCC 555
 Qy 421 ATGTTGCCGACCTGTTGCTTCTACCCCGGCTGACCGGTCATGAGCTGTGCTG 480
 Db 556 ATGTTGCCGACCTGTTGCTTCTACCCCGGCTGACCGGTCATGAGCTGTGCTG 615
 Qy 481 CGGTGAGCTTCTATGGTGGCTTGGAGGATGGACTCTGTCTTACACCGCCCTGTG 540
 Db 616 CGGTGAGCTTCTATGGTGGCTTGGAGGATGGACTCTGTCTTACACCGCCCTGTG 675
 Qy 541 GGGCAGACAATGTTATCTAGAGCCTGTACCTCAACGACTCCACTATGACGGACAT 600
 Db 676 GGGCAGACAATGTTATCTAGAGCCTGTACCTCAACGACTCCACTATGACGGACAT 735
 Qy 601 ACCGTGGCGGACTGACATGAGGCTGGCCAGCTGGCAGTGGAGTGTGGTGGGCTG 660
 Db 736 ACCGTGGCGGACTGACATGAGGCTGGCCAGCTGGCAGTGGAGTGTGGTGGGCTG 795
 Qy 661 GATGACTTTAGAGAGTACAGAGCTGGGCTGGCCAGCTATGACTATGTGGAGTG 720
 Db 796 GATGACTTTAGAGAGTACAGAGCTGGGCTGGCCAGCTATGACTATGTGGAGTG 855
 Qy 721 AGCAACACAGCTTCTCAGTGGCTATGTGAGATGAGATTTGATTTGACCGGCTAGG 780
 Db 856 AGCAACACAGCTTCTCAGTGGCTATGTGAGATGAGATTTGATTTGACCGGCTAGG 915
 Qy 781 GCCTTCCAGGCTATGAGGCTCCACTGTAAACATGACACGCTGGAGCCGCTGCTCCT 840
 Db 916 GCCTTCCAGGCTATGAGGCTCCACTGTAAACATGACACGCTGGAGCCGCTGCTCCT 975
 Qy 841 GCGGGGTGGAATGTCGCTTCCGGGCTGGCCCTGCCATGGCTGGGAGGGGAGCCCATG 900
 Db 976 GCGGGGTGGAATGTCGCTTCCGGGCTGGCCCTGCCATGGCTGGGAGGGAGCCCATG 1035
 Qy 901 CCCACACACTAGGGGGCAACCTGGGGAGCCCGAGAGCCCGGCTGTCTAGTCCCGCTT 960
 Db 1036 CCCACACACTAGGGGGCAACCTGGGGAGCCCGAGAGCCCGGCTGTCTAGTCCCGCTT 1095
 Qy 961 GCGGCGCTGTGGCTGCTTCTGAGTGGCGCTTCCCTTTTGGGGGCTGCTGTACTC 1020
 Db 1096 GCGGCGCTGTGGCTGCTTCTGAGTGGCGCTTCCCTTTTGGGGGCTGCTGTACTC 1155
 Qy 1021 TTCAGCGAAATCTCCTTCTATGATGGTGAACAATTCCTTCGGGCACTGGGAGGC 1080
 Db 1156 TTCAGCGAAATCTCCTTCTATGATGGTGAACAATTCCTTCGGGCACTGGGAGGC 1215
 Qy 1081 ACCTTCCGCGACGCCCTTGGTGGCGGCTGGCCAGCTCCACACACTTCAGCAGCTTG 1140
 Db 1216 ACCTTCCGCGACGCCCTTGGTGGCGGCTGGCCAGCTCCACACACTTCAGCAGCTTG 1275
 Qy 1141 GAGCTGGAGCCAGAGGCGACAGCCCTGGCCAGCCGAGGGGAGCCGACCGCC 1197

Db 1276 GAGCTGGAGCCAGAGCCAGCAGCCCTGGCCAGCCCGAGGGAGCCGACCGCC 1332

RESULT 2
 US-08-445-640-7
 : Sequence 7, Application US/08445640
 : Patent No. 5709858
 : GENERAL INFORMATION:
 : APPLICANT: Godowski, Paul J.
 : APPLICANT: Mark, Melanie R.
 : APPLICANT: Scadden, David T.
 : APPLICANT: Baker, Kevin P.
 : APPLICANT: Baron, Will F.
 : TITLE OF INVENTION: Protein Tyrosine Kinases
 : NUMBER OF SEQUENCES: 35
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Genentech, Inc.
 : STREET: 460 Point San Bruno Blvd
 : CITY: South San Francisco
 : STATE: California
 : COUNTRY: USA
 : ZIP: 94080
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: 5.25 inch, 360 kb floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: patin (Genentech)
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/445,640
 : FILING DATE: 22-MAY-1995
 : CLASSIFICATION: 435
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 08/170558
 : FILING DATE: 20-DEC-1993
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 08/157563
 : FILING DATE: 23-NOV-1993
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Hasak, Janet E.
 : REGISTRATION NUMBER: 28,616
 : REFERENCE/DOCKET NUMBER: 854C2
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 415/225-1896
 : TELEFAX: 415/952-9881
 : TELEX: 910/371-7168
 : INFORMATION FOR SEQ ID NO: 7:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 1197 bases
 : TYPE: nucleic acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : US-08-445-640-7

Query Match 100.0%; Score 1197; DB 2; Length 1197;
 Best Local Similarity 100.0%; Pred. No. 2e-287;
 Matches 1197; Conservative 0; Mismatches 0; Indels 0; Gaps 0:

Qy 1 GATGCTGACATCAAGGACATTTTGATCTGCAAGTCCGCTATGCCCTGGGCATGCAG 60
 Db 1 GATGCTGACATCAAGGACATTTTGATCTGCAAGTCCGCTATGCCCTGGGCATGCAG 60

Qy 61 GACCGGACCATCCAGACAGTGCATCTCTCCAGCTCTCTGTCTCAGATTCACACTGCC 120
 Db 61 GACCGGACCATCCAGACAGTGCATCTCTCCAGCTCTCTGTCTCAGATTCACACTGCC 120

Qy 121 GCGCGCCACAGAGTTGGAGAGCAGTACGGGATGGGCTGGTGGCCCGCAGGTCG 180
 Db 121 GCGCGCCACAGAGTTGGAGAGCAGTACGGGATGGGCTGGTGGCCCGCAGGTCG 180

Qy 181 GTGTTTCCCAAGGAGGAGGAGTACTTTCAGGTGGATCTACACAGCTGCACCTGGTGGCT 240
 Db 181 GTGTTTCCCAAGGAGGAGGAGTACTTTCAGGTGGATCTACACAGCTGCACCTGGTGGCT 240

QY 241 CTGGTGGGACACGAGGAGCGCATGCGGGGGGCTGGGCAAGAGTTCTCCGGAGCTAC 300
DB 241 CTGGTGGGACACGAGGAGCGCATGCGGGGGGCTGGGCAAGAGTTCTCCGGAGCTAC 300
QY 301 CGGCTGGTACTCCCGGAGTGTCCCGCTGGATGGGCTGGAAGGACCGCTGGGGTAC 360
DB 301 CGGCTGGTACTCCCGGAGTGTCCCGCTGGATGGGCTGGAAGGACCGCTGGGGTAC 360
QY 361 GAGGTGATCTCAGGCAATGAGGACCTGAGGGAGTGTCTGAAGGACCTTGGGCCCCC 420
DB 361 GAGGTGATCTCAGGCAATGAGGACCTGAGGGAGTGTCTGAAGGACCTTGGGCCCCC 420
QY 421 ATGGTTGCCGAGTGTCTCTTACCCCGGGCTGACCGGGTCAAGAGCTGTGTCTG 480
DB 421 ATGGTTGCCGAGTGTCTCTTACCCCGGGCTGACCGGGTCAAGAGCTGTGTCTG 480
QY 481 CGGCTAGAGCTCATGGCTGCTCTGAGGAGTGTCTCTTACACCGCCCTGTG 540
DB 481 CGGCTAGAGCTCATGGCTGCTCTGAGGAGTGTCTCTTACACCGCCCTGTG 540
QY 541 GGCAGACAATGTATTTATCTAGGCGGTGTACCTCAACGACTCCACCTATGACGACAT 600
DB 541 GGCAGACAATGTATTTATCTAGGCGGTGTACCTCAACGACTCCACCTATGACGACAT 600
QY 601 ACCGTGGGCGGACTGAGTATGGGGTCTGGGCGAGCTGGCAGATGGTGTGGGGTGG 660
DB 601 ACCGTGGGCGGACTGAGTATGGGGTCTGGGCGAGCTGGCAGATGGTGTGGGGTGG 660
QY 661 GATGACTTTAGGAAGAGTCAGAGCTGCGGGTCTGGCCAGGCTATGACTATGGGATGG 720
DB 661 GATGACTTTAGGAAGAGTCAGAGCTGCGGGTCTGGCCAGGCTATGACTATGGGATGG 720
QY 721 AGCAACACAGCTTCTCCAGTGGCTATGTGGAGTGGAGTTTGAAGTTGACCGGCTGAGG 780
DB 721 AGCAACACAGCTTCTCCAGTGGCTATGTGGAGTGGAGTTTGAAGTTGACCGGCTGAGG 780
QY 781 GCCTTCAGGCTATGAGGTCTCACTGTAAACATGACACGCTGGGAGCCGCTGTGCT 840
DB 781 GCCTTCAGGCTATGAGGTCTCACTGTAAACATGACACGCTGGGAGCCGCTGTGCT 840
QY 841 GGGGGGTGGAATGTGCTTCCGGCTGCGCTGCGCATGGCTGGGAGGGGAGCCCATG 900
DB 841 GGGGGGTGGAATGTGCTTCCGGCTGCGCTGCGCATGGCTGGGAGGGGAGCCCATG 900
QY 901 CGCCACAACCTAGGGGGCAACCTGGGGGACCCAGAGCCCGGCTGTCTCAGTGGCCCTT 960
DB 901 CGCCACAACCTAGGGGGCAACCTGGGGGACCCAGAGCCCGGCTGTCTCAGTGGCCCTT 960
QY 961 GGGGGCGTGTGGCTGCTTCTGAGTCCCGCTTCTTTGGGGGCGCTGTGTTACTC 1020
DB 961 GGGGGCGTGTGGCTGCTTCTGAGTCCCGCTTCTTTGGGGGCGCTGTGTTACTC 1020
QY 1021 TTCCAGGCAATCTCTTCACTCTCATGTGGTGAACAATTCCTCTCCGGCACTGGAGGC 1080
DB 1021 TTCCAGGCAATCTCTTCACTCTCATGTGGTGAACAATTCCTCTCCGGCACTGGAGGC 1080
QY 1081 ACTTCCCGCCAGCCCTTGTGGCGCTGCGCCCTGCGCCACCTCCCAACCACTTCAGAGCTTG 1140
DB 1081 ACTTCCCGCCAGCCCTTGTGGCGCTGCGCCCTGCGCCACCTCCCAACCACTTCAGAGCTTG 1140
QY 1141 GAGCTGGAGCCAGAGGCGGAGCCGCTGGCGGCTGCGCCAGCGGAGGCGCCGAGCCG 1197
DB 1141 GAGCTGGAGCCAGAGGCGGAGCCGCTGGCGGCTGCGCCAGCGGAGGCGCCGAGCCG 1197

RESULT 3

US-08-336-343A-1

: Sequence 1, Application US/08336343A

: Patent No. 567144

: GENERAL INFORMATION:

: APPLICANT: Ullrich, Axel

: APPLICANT: Alves, Frauke

; TITLE OF INVENTION: CCK-2, A No. 567714461 Receptor Tyrosine Kinase
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/336,343A
; FILING DATE: 08-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7683-065
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3962 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 321..3077
; US-08-336-343A-1

Query Match 99.8%; Score 1192.2; DB 2; Length 3962;
Best Local Similarity 99.7%; Pred. No. 4.3e-286;
Matches 1194; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GATGCTGACATGAGGAGGACATTTTATCTGCTCCGCAAGTGGCGCTATGCCCTGGGCATGCAG 60
DB 375 GATGCTGACATGAGGAGGACATTTTATCTGCTCCGCAAGTGGCGCTATGCCCTGGGCATGCAG 434
QY 61 GACCGGACCATCCAGACAGTGCATCTCTGCTTCCAGCTCCTGGTTCAGATTCCACTGCC 120
DB 435 GACCGGACCATCCAGACAGTGCATCTCTGCTTCCAGCTCCTGGTTCAGATTCCACTGCC 494
QY 121 GCCCGCCACAGAGTTGGAGACAGTGCAGGGGATGGGGCTGGTGGCCCGCAGGCTCG 180
DB 495 GCCCGCCACAGAGTTGGAGACAGTGCAGGGGATGGGGCTGGTGGCCCGCAGGCTCG 554
QY 181 GTGTTTCCCAAGGAGGAGGAGTACTTCCAGGTGGATCTACAAGCTGCACCTTGGTGGCT 240
DB 555 GTGTTTCCCAAGGAGGAGGAGTACTTCCAGGTGGATCTACAAGCTGCACCTTGGTGGCT 614
QY 241 CTGTTGGGCAACCCAGGAGCGCATGCCGGGCTGGGCAAGGAGTTCTCCCGGAGCTAC 300
DB 615 CTGTTGGGCAACCCAGGAGCGCATGCCGGGCTGGGCAAGGAGTTCTCCCGGAGCTAC 674
QY 301 CGGCTGGCTTACTCCCGGGATGTCGCCGCTGGATGGGTGAAGACCGCTGGGGTACAG 360
DB 675 CGGCTGGCTTACTCCCGGGATGTCGCCGCTGGATGGGTGAAGACCGCTGGGGTACAG 734
QY 361 GAGGTGATCTCAGGCAATGAGGACCTGAGGGAGTGGTGTGAGGACCTTGGGCCCCCCC 420
DB 735 GAGGTGATCTCAGGCAATGAGGACCTGAGGGAGTGGTGTGAGGACCTTGGGCCCCCCC 794

QY 421 ATGGTGGCCGACATGGTTCGGTCTTACCCCGGGCTGACCGGGTCATGACGGTCTGTCTG 480
DB 795 ATGGTGGCCGACATGGTTCGGTCTTACCCCGGGCTGACCGGGTCATGAGTCTGTCTG 854
QY 481 CGGTAGAGCTCTATGGTGGCTCTGGAGGATGAGTCTCTGTCTTACACCGCCCTGTG 540
DB 855 CGGTAGAGCTCTATGGTGGCTCTGGAGGATGAGTCTCTGTCTTACACCGCCCTGTG 914
QY 541 GGGCAGACAATGATTTATCTAGGCGGTGTACCTAACGACTCCACTATGACGGACAT 600
DB 915 GGGCAGACAATGATTTATCTAGGCGGTGTACCTAACGACTCCACTATGACGGACAT 974
QY 601 ACCGTGGCGGACATGACGATATGGGGTCTGGCCAGCTGGCAGATGTGTGGGGCTG 660
DB 975 ACCGTGGCGGACATGACGATATGGGGTCTGGCCAGCTGGCAGATGTGTGGGGCTG 1034
QY 661 GATGACTTTAGGAAGTACGAGTCTGGGCTCTGGCCAGGCTATGACTATGTGGGATG 720
DB 1035 GATGACTTTAGGAAGTACGAGTCTGGGCTCTGGCCAGGCTATGACTATGTGGGATG 1094
QY 721 AGCAACACAGCTCTCCAGTGGCTATGTGGAGATGAGTTTGAGTTTGACCGGCTGAG 780
DB 1095 AGCAACACAGCTCTCCAGTGGCTATGTGGAGATGAGTTTGAGTTTGACCGGCTGAG 1154
QY 781 GCGTTCCAGGCTATGACGATCCACTGTAAACAATGACACGCTGGGAGCCGCTGCTCCT 840
DB 1155 GCGTTCCAGGCTATGACGATCCACTGTAAACAATGACACGCTGGGAGCCGCTGCTCCT 1214
QY 841 GCGGGGTGGAATGTGCTTCCGGGTGGCCCTGCCATGGCTGGGAGGGAGCCCATG 900
DB 1215 GCGGGGTGGAATGTGCTTCCGGGTGGCCCTGCCATGGCTGGGAGGGAGCCCATG 1274
QY 901 CGCCACAACCTAGGGGGAACCTGGGGGACCCAGAGCCCGGGCTGTCTCAGTGCCTT 960
DB 1275 CGCCACAACCTAGGGGGAACCTGGGGGACCCAGAGCCCGGGCTGTCTCAGTGCCTT 1334
QY 961 GCGGCGCTGTGGCTCGCTTTCTGAGTGGCGCTTCTCTTTGGGGGCGCTGTGTTACT 1020
DB 1335 GCGGCGCTGTGGCTCGCTTTCTGAGTGGCGCTTCTCTTTGGGGGCGCTGTGTTACT 1394
QY 1021 TTCACGGAATCTCCTTCTATCTGTGTGTAACAAATCTCTCCGGGACTGGGAGG 1080
DB 1395 TTCACGGAATCTCCTTCTATCTGTGTGTAACAAATCTCTCCGGGACTGGGAGG 1454
QY 1081 ACCTTCCGCGGACCCCTGTGTGGCGGCTGGCCCACTCCCAACAACTTACAGAGTTG 1140
DB 1455 ACCTTCCGCGGACCCCTGTGTGGCGGCTGGCCCACTCCCAACAACTTACAGAGTTG 1514
QY 1141 GAGCTGGAGCCGAGGCGGACGAGCCGCTGGCCAGCCCGGAGGGAGCCGACCGCC 1197
DB 1515 GAGCTGGAGCCGAGGCGGACGAGCCGCTGGCCAGCCCGGAGGGAGCCGACCGCC 1571

RESULT 4

US-08-336-343A-3
; Sequence 3, Application US/08336343A
; Patent No. 5677144
; GENERAL INFORMATION:
; APPLICANT: Ulrich, Axel
; APPLICANT: Alves, Frauke
; TITLE OF INVENTION: CCK-2, A No. 5677144el Receptor Tyrosine Kinase
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESS: Pennie & Edmonds
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/336.343A
; FILING DATE: 08-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7683-065
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3157 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 370..2934
; US-08-336-343A-3

Query Match 27.4%; Score 327.4; DB 2: Length 3157;

Best Local Similarity 58.8%; Pred. No. 2.2e-72;

Matches 644; Conservative 0; Mismatches 436; Indels 15; Gaps 4;

QY 13 AAGGACATTTTATCTCTGCAAGTGGCGCTATGCCCTGGGCATGACGAGCCGACCATC 72
DB 433 AAGCTCAGGTTAATCCAGCTATATGCGCTATCTCTGGGCATGTCAGAGGCCAGATT 492
QY 73 CCAGACAGTGCATCTCTGCTTCCAGCTCCCTGCTCAGATTCACCTGCGCCGCCACAGC 132
DB 493 CCAGATGAGGACATCACAGCTTCCAGTCACTGGTTCAGAGTCCACAGCTGCCAAATATGA 552
QY 133 AGTTTGGAGACAGTGAACGGGATGGGGCTGCTGCTGCCCGAGGCTGCTTCCCAAG 192
DB 553 AGCTGGACTCAGAGAGAGGAGTGGAGCTGTGCTGCTGAGATTCAGTGGAACTGAT 612
QY 193 GA---GGAGGAGTACTTGCAGTGGATGTACAAACAGCTGCACCTGGTGGCTGTGGGCG 249
DB 613 GACCTGAAGGAGTCTTCTGCAGATTCAGTTCACACCCCTCCATTTCATCCTCTGTGGGG 672
QY 250 ACCCAGGAGCGATGCGCGGCGCTGGGCAAGGAGTTCCTCCCGAGGCTACCGGCTGGCT 309
DB 673 ACCCAGGCGGCATGTCAGGAGGTCATGGCATCGAGTTTGGCCCCCATGTACAGATCAAT 732
QY 310 TACTCCCGGATGGTGGCGCTGGATGGGCTGGAAGGACCGCTGGGGTTCAGGAGTGCATC 369
DB 733 TACAGTCGGGATGGGACTGCTGTGATCTCTTGGCGGACCGCTCATGGGAACAGGTGCTG 792
QY 370 TCAGGCAATGAGGACCTTGAAGGAGTGTGTGAAGGACCTTGGGCCCCCCTATGTTGCC 429
DB 793 GATGAAATAGTAAACCCCTATGACATTTCTCTAAAGGACTTGGAGCCGCCATTTAGGCC 852
QY 430 CGACTGTTTCGGTCTTACCCCGGCTGACCGGCTGACGAGTGTCTGCTCGGGTAGAG 489
DB 853 AGATTTGTCGGGTTTCATTCACCGGACCTCCAGTGAATGTGTGTATGAGATGGAG 912
QY 490 CTCTATGGCTGCTCTGGAGGATGAGTCTCTGTCTTACACCGCCCTGTGGGCGAGACA 549
DB 913 CTTACGGCTGTCTGGCTAGATGCTTGTGTCTTACAACTGCTCCAGTCCGAGCAG 972
QY 550 ATGATTTATCTGAGGC-----CGTGTACCTCAAGGACTCCACCTATGACGAGCATACC 603
DB 973 TTTGTACTCCCTGGAGGTTTCCATCATTTATCTGAATGATTTCTGTCTATGATGAGCTGT 1032
QY 604 GTGGCGGACTGCAGTATGGGGTCTGGGCCAGCTGGGAGATGGTGTGGTGGGCTGGAT 663

Db 1033 GGATACAGCATGACAGAAGG---CTAGGCCAATGACCGATGGTGTCTGGCCCTGGAC 1089
Qy 664 GACTTTAGGAAGAGTCAGGAGCTCGGGTCTGGCCAGGCTATGACTATGTGGGATGAGC 723
Db 1090 GATTTCACCCAGACCATGAATACCAGCTGTGGCCGGCTATGACTATGTGGGCTGGCG 1149
Qy 724 AACACAGCTTCTCCAGTGCTATGTGGAGATGGAGTTTGAAGTTTGAACGGCTGAGGGCC 783
Db 1150 AACAGAGTCCACCAATGGCTACATGAGATCATGTTTGAATTTGACCGCATCAGGAAT 1209
Qy 784 TTCCAGCTATGAGGCTCCACTGTACAAACATGCACACGCTGGGAGCCGCTCTGCCGTC 843
Db 1210 TTCACTACCATGAGGCTCCACTGTCAACACATGTTTGTAAAGGTGTGAAGATCTTTAAG 1269
Qy 844 GGGGTGAATGTGCTTCCGGCTGGCCCTGCGCATGGCTGGAGGGGAGCGCATGCGC 903
Db 1270 GAGTACAGTGTACTTCCGCTCTGAA---GCCAGTGTGGGAACCTATGCCATTTCC 1326
Qy 904 CACAACTAGGGGCAACCTGGGGGACCCAGAGCCCGGGCTCTCAGTGGCCCTTGGC 963
Db 1327 TTCCGCTTCTCGTGTGATGACGTCAACCCAGTGTCTGGTGTGTCACGGTGTCTCTCCAC 1386
Qy 964 GGGCTGTGCTGCTTCTCTGAGTCCGCTTCTCTTTGCGGGGCGCTTACTCTTC 1023
Db 1387 CACGAAATGGCAGTGGCCATCAAGTGTCAATACCATTTTGCAGATACCTGGATGTTC 1446
Qy 1024 AGGGAATCTCTTCTATCTCTGATGTGTGGAACAATTCCTCTCCGGCACTGGGAGCAC 1083
Db 1447 AGTGATACCTTCCATCAGATGTCGAATGTACAAACACTCTGAAGCCCTGCCACC 1506
Qy 1084 TTCCGCGCCAGCCCC 1098
Db 1507 TCTCCTATGGCACCC 1521

RESULT 5

US-08-336-343A-5/c
; Sequence 5, Application US/08336343A
; Patent No. 5677144
; GENERAL INFORMATION:
; APPLICANT: Ullrich, Axel
; APPLICANT: Alves, Frauke
; TITLE OF INVENTION: CK-2, A No. 5677144el Receptor Tyrosine Kinase
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/336, 343A
; FILING DATE: 08-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7683-065
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3157 base pairs
; TYPE: nucleic acid

; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-336-343A-5

Query Match 27.4%; Score 327.4; DB 2; Length 3157;
Best Local Similarity 58.8%; Pred. No. 2.2e-72;
Matches 644; Conservative 0; Mismatches 436; Indels 15; Gaps 4;

Qy 13 AAGGACATTTTGTATCTTCCCAAGTGGCGCTATGCTTCCCTGGGCACTGAGGACCGGACCATC 72
Db 2725 AAAGCTCAGGTTAATCCAGCTATATGCGCTATCTCTGGGCACTGTCAGGAGGCCAGATT 2666
Qy 73 CCAGACAGTACATCTCTGCTTCCAGCTCTCTGGTTCAGATTCACATCCAGCTGCGGCCGCCAGC 132
Db 2665 CCAGATGAGGACATCACAGCTTCCAGTCACTGAGTGGTCCAGATCCACAGCTGCCAAATATGA 2606
Qy 133 AGTTGGAGAGCAGTGACGGGATGGGCGCTGGTGGCCCGCAGGCTGGTGTCTTCCCAAG 192
Db 2605 AGGCTGGACTCAGAAAGAGGATGGAGCTTGGTGGCCCTGAGATTCAGTGGAACTGAT 2546
Qy 193 GA---GGAGGAGTACTTTCAGTGGATCTACAACTGACACTGACCTTGGTGGTCTCTGGTGGC 249
Db 2545 GACCTGAAGGAGTTTCTGCAGATTGACTTGCACACCTCCATTTTATCACTCTCTGGTGGG 2486
Qy 250 ACCAGGACCGCATCGCGGGGCTGGGGAAGAGTTCTCCCGGAGCTACCGGCTCGCT 309
Db 2485 ACCAGGGGCCCATGTCAGAGGCTCATGGCATCGAGTTTGGCCCATGTACAAGATCAAT 2426
Qy 310 TACTCCCGGATGTGCGCGCTGGATGGGTGGTGAAGACCGCTGGGTCAGGAGGTGATC 369
Db 2425 TACAGTGGGATGGACATCGCTGGATCTTGGCGGAACCTCATGGAACAGAGTGTG 2366
Qy 370 TCAGCAATAGGACCTCTAGGAGTGTGTCTGAAGACCTTGGGCGCCCATCTGTTGCC 429
Db 2365 GATGAAATAGTAACCCCTATGACATTTCTTAAGACTTGGAGCCGCCCATTTGTAGCC 2306
Qy 430 CGACTGGTTCCTTCTACCCCGGCTGACCGGGTCTATGAGCGTCTGTCTCGGGTGTAG 489
Db 2305 AGATTGTCGGTTCATTCACAGTCCAGCCAGCTCCCATGAATGTGTGTATGAGAGTGGAG 2246
Qy 490 CTCATGCTCCCTCTGGAGGATGGATCTCTGTCTTACACCGCCCTGTGGGCGACACA 549
Db 2245 CTTACGCTGTGTCTGGCTAGTGGTGTGTGTCTTACAACTTCCAGTCCAGCTGGCAGAG 2186
Qy 550 ATGTATTTATCTGAGGCTCTTACCTCAACGACTCCACCTATGACGGACATACC 603
Db 2185 TTTGTACTCCCTGGAGGTTCCATCATTTATCTGAATGATTTCTATGATGGAGCTGTT 2126
Qy 604 GTGGCGGACTGCAAGTATGGGGGCTCTGGGCGAGCTGGCAGATGTGTGTGGGCTGGAT 663
Db 2125 GGATACAGCATGACAGAAGG---CTAGGCCAATTTGACCGATGTGTCTGTGGCTGGAC 2059
Qy 664 GACTTTAGGAAGAGTCTGAGGCTCGGGTCTGGCCAGCTATGACTATGTGGGATGAGC 723
Db 2058 GATTTCACCCAGACCCATGAATACAGTGTGGCCCGGCTATGACTATGTGGGCTGGCGG 2009
Qy 724 AACCAAGCTTCTCCAGTGGCTATGTGGAGTGGAGTTTGAAGTTTGAACCGCTGAGGGCC 783
Db 2008 AACGAGATGCCACCATGGCTACATTCAGATCATGTTTGAATTTGACCGCATCAGGAAT 1949
Qy 784 TTCAGGCTATGAGGTCCTGTAACAACATGCACAGCTGGGAGCCCGCTCTCCCTGGC 843
Db 1948 TTCCTACCATGAAGTCCACTGTCAACAACATGTTTGTAAAGTGTGAAGATCTTTAAG 1889
Qy 844 GGGTGGATGCTCTCCCGGCTGGCCCTGCCATGGCTGGGAGGGGAGCCCATCGC 903
Db 1888 GAGGTACAGTGTCTTCCGCTCTGAA---GCCAGTGTGGGGAACCTAATGCAATTTCC 1832
Qy 904 CACAACCTAGGGGCAACCTGGGGGACCCAGAGCCCGGGCTGTCTCAGTGGCCCTTGGC 963

Db 1831 TTCCCGCTTGCTGGATGACGTCAACCCAGTCTCGGTTTGTACGGGTGCTCTCCAC 1772
Qy 964 GCGCGTGTGGCTCGCTTTCTTCAGTGGCGCTTCCTTTTGGGGGCGCTGGTACTTTC 1023
Db 1771 CACCGAATGGCCAGTGCATCAAGTCAATACCATTTTGCAGATACCTGGATGTTTC 1712
Qy 1024 AGCGAATCTCCTTCATCTCTGATGTGGTGAACAATTCCTCTCCGGCACTGGGAGGCACC 1083
Db 1711 AGTGAGATCACTTCCCAATCAGATGCTGCAATGTACAACAACACTGGAAGCCCTGCCACC 1652
Qy 1084 TTCCCGCCAGCCGCC 1098
Db 1651 TCTCTATGGCACCC 1637

RESULT 6

US-08-456-647B-19
; Sequence 19, Application US/08456647B
; Patent No. 5811516
; GENERAL INFORMATION:
; APPLICANT: Lemke Ph.D. et al., Greg E.
; TITLE OF INVENTION: PROTEIN-TYROSINE KINASE GENES
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/456.647B
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/237,401
; FILING DATE: 02-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/884,486
; FILING DATE: 15-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Wetherell Ph.D., John R.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: 07251/007002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 678-5070
; TELEFAX: (619) 678-5099
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3120 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; IMMEDIATE SOURCE:
; CLONE: Tyro-10
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 485..3047
US-08-456-647B-19

Query Match 26.8%; Score 321; DB 3; Length 3120;
Best Local Similarity 58.4%; Pred. No. 8.2e-71;
Matches 640; Conservative 0; Mismatches 440; Indels 15; Gaps 4;

Qy 13 AAGGACATTTTGATCTCGCAAGTGGCGCTATGCCCTTGGGCATGACGAGCCGACCATC 72

Db 548 AAAGCTCAGGTTAAATCCAGCCATATGCGGCTATCTCTGGGCATGTCTAGAGGCGCACATT 607
Qy 73 CCAGCAGATGACATCTCTGCTTCCAGCTCCTGGTCCAGATTCACATGCGCGCGCCACAGC 132
Db 608 CCAGATGAGGACATCACAGCCTCAAGTCAGTGGTCAGAAATCCAGGCTGCCAATATGGG 667
Qy 133 AGGTTGGAGAGCAGTGTACGGGAGTGGGCGTGGTGGCCCGGAGGGTCCGTTCTTTCCCAAG 192
Db 668 AGGCTGGACTCTGAAGAAGGAGATGGAGCTGGTCTCTGAGATTCAGATTCAGACCCGAT 727
Qy 193 GA---GGAGGAGTACTTTCAGGTGATCTACAACGACATGCACCTGGTGGCTCTGGGCG 249
Db 728 GACCTGAAGGAATTTCTGCAGATTGACTTGGCAACCTTACACTTTATCACTCTTGTGGGG 787
Qy 250 ACCCAGGACGATGCGCGGGGCTGGGCAAGGAGTTCTCCCGGAGCTACCGGTGGCT 309
Db 788 ACCCAGGCGGCATGACAGGGGTCTATGGCATTGAATTTGACCCATGTACAAGATCAAC 847
Qy 310 TACTCCCGGATGGTGGCGCTGGATGGGCTGGAAGACCGCTGGGGTTCAGAGGTGATC 369
Db 848 TACAGTGGGATGGCAGTGGCTGGATCTCTGGCGTAACCGGCATGGAAGCAGGTGCTT 907
Qy 370 TCAGCAATGAGGACCCCTGAGGAGTGGTGTGAAGGACCTTGGCCCCCATGCTTGGCC 429
Db 908 GATGAAACAGTAACCCCTTATGATGTATTCCTGAAGGACTTGGAGCCACCATGCTGCC 967
Qy 430 CGACTGGTTTCGCTTCTACCCCGGCTGACGGGTCATGAGGCTCTGTCTCGGGTAGAG 489
Db 968 AGATTTGTTCGCTTATCCAGTCACTGACCACTCCATGAAGCTGTGCATGAGGTTGAG 1027
Qy 490 CTCTATGGCTCTGGAGGATGGACTCTCTGTCTTACACCGCCCTGTGGGCGACAGA 549
Db 1028 CTTTATGGTTGTCTGGCTAGATGGCTTGGTATCTCTACAATGTCTCCAGTCCAGCAG 1087
Qy 550 ATGTATTTATCTGAGC-----CGTGTACCTCAAGACTCCACCTATGAGGACATACC 603
Db 1088 TTTGTACTCCCTGGAGGCTCCATCATTTATCTGAATGATTCTGTCTATGATGGAGCTTT 1147
Qy 604 GTGGCGGACTGCAGTATGGGGGTCTGGCCAGCTGGCAGATGGTGGTGGGCTGGAT 663
Db 1148 GGTACAGCATGACTGAAGG---CTAGCCAGTTGACTGATGAGATATCGGGCTGGAT 1204
Qy 664 GACTTTAGGAAGATCAGGAGCTCGGGTCTGGCCAGGCTATGACTATGTGGATGAGC 723
Db 1205 GATTTTACCAGACCCATGAATACCACGCTGGCTGGCTGCTGACTACGTGGATGGCGG 1264
Qy 724 AACCACAGCTTCTCCAGTGGCTATGTGAGATGGAGTTTGAGTTTGACCGGCTGAGGSCC 783
Db 1265 AACGAAAGTGTACCAACGGTTTCATTGAGATCATGTTTGAATTTGACCGAATCAGGAAT 1324
Qy 784 TTCCAGGCTATGCAGTCCACTGTAAACAACATGCACACGCTGGGAGCCCGCTGTGCTGGC 843
Db 1325 TTTACTACCATGAAGTCCACTGCAACAACATGTTTGTGAAGGTGTGAAGATTTTAAAG 1384
Qy 844 GGGGTGGAATGTGCTTCCGGCTGGCCCTGCCATGGCTGGGAGGGGAGGCCATCGCC 903
Db 1385 GAGGTCCAGTGTACTTTTCGCTCGGA---AGCCAGCAGTGGGAACCCACCTGCTCTAC 1441
Qy 904 CACAACCTAGGGGGCAACCTGGGGGACCCAGAGCCCGGGCTGTCTCAGTGGCCCTTGGC 963
Db 1442 TTTCCCTTGGTCTTGGACGATGTGAACCCCAAGTGGCCCGTTTGTACCGGTGCCCTCCAC 1501
Qy 964 GCGCGTGTGGCTCTTCTGCAGTGGCGCTTCCCTTCTTTGGGGGCGCTGGTTACTCTTC 1023
Db 1502 CACCGAATGGCCAGTGCATCAAGTGCACCAATACCATTTTGGCAGACAGTGGATGATTC 1561
Qy 1024 AGCGAATCTCCTTCACTCTCTGATGTGGTGAACAATTCCTCTCCGGCACTGGGAGGCACC 1083
Db 1562 AGCGAGATCACTTTCCAAATCAGATGCTGCAATGTATAACAACCTCTGGAGCCCTTCCCACC 1621
Qy 1084 TTCCCGCCAGCCGCC 1098

COUNTRY: USA
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/670,707A
FILING DATE: 26-JUN-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US94/13200
FILING DATE: 15-NOV-1994
NUMBER OF SEQUENCES: 40
APPLICATION DATA:
APPLICATION NUMBER: US 08/212,133
FILING DATE: 11-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/864,004
FILING DATE: 07-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Greenlee, Lorraine L.
REGISTRATION NUMBER: 27,894
REFERENCE/DOCKET NUMBER: 75-95F
TELEPHONE: 303/499-8080
TELEFAX: 303/499-8089
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 6402 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: not relevant
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Pig
FEATURE:
NAME/KEY: CDS
LOCATION: 1..6402
US-08-670-707A-36

Query Match 3.9%; Score 46.6; DB 3; Length 6402;
Best Local Similarity 45.9%; Pred. No. 0.0066;
Matches 197; Conservative 0; Mismatches 229; Indels 3; Gaps 1;
QY 22 TTGTATCTCTCCCAAGTGGCGCTATGCCCTGGCATGCGACCGGACCATCCACAGT 81
DB 5908 TGTGATTAAACAGTTGCAGCATGCCCTGGGAATGCAGATAAAGCGATATCAGACTCA 5967
QY 82 GACATCTCTGTCTCCAGTCTCTGGTTCAGATTCCACTGGCGCGCCGACAGAGTTGGAG 141
DB 5968 CAGATCAGCGCTCTCTCCACCATAGCATATATTGGCCACTGTCTCTCCACAGCC 6027
QY 142 AGCAGTGACGGGATGGCGCTGGTGGCCCGCAGGCTGGTGTTCCTCCCAAGGAGGAGGAG 201
DB 6028 CGACTTCACCTCCAGGGCGGACGAATGCTGGCGACCCCGGTGAGCAGCGCAGAGGAG 6087
QY 202 TACTTCAGGTGGATACAGCACTGCACCTGGTGGTCTCTGGTGGGACCCAGGAGCGG 261
DB 6088 TGGCTGAGGTGGACCTGCAGAGAGCGGTGAAGGTCACAGGCATCACCACCCAGGGCGGTG 6147
QY 262 CATGCCGGGGCGCTGGCAAGGAGTTCTCCCGGAGTACCGGCTGGCTTACTCCCGGGAT 321
DB 6148 AGTCCCTGCTCAGCAGCATGTATGTGAGAGGATTCCTCTGTCTCAGT---AGTCAGGAC 6204
QY 322 GGTGCGCGGTGGATGGGTGGAGAGACCGCTGGGTTCAGAGGTGATCTCAGGCAATGAG 381
DB 6205 GGCGCGCGGTGGACCTGTGTCTTCAGGACCGCCACACGAAGGTTTTTCAGGGCAATCAG 6264
QY 382 GACCTTGAGGAGTGGTGGTGGAGACCTTGGCGCCCGCCCATGTTGCCCGACTGTTGCG 441

DB 6265 GACTCTCTCCACCCCGTGGTGAACGCTCTGGACCCCGCTGTTCCACGCGCTACTGAGG 6324
QY 442 TTCTACCC 450
DB 6325 ATCCACCC 6333
RESULT 9
US-08-670-707A-38
Sequence 38, Application US/08670707A
Patent No. 5859204
GENERAL INFORMATION:
APPLICANT: Lollar, John S.
TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: USA
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/670,707A
FILING DATE: 26-JUN-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US94/13200
FILING DATE: 15-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/212,133
FILING DATE: 11-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/864,004
FILING DATE: 07-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Greenlee, Lorraine L.
REGISTRATION NUMBER: 27,894
REFERENCE/DOCKET NUMBER: 75-95F
TELEPHONE: 303/499-8080
TELEFAX: 303/499-8089
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 4334 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: not relevant
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Factor VIII lacking B domain
FEATURE:
NAME/KEY: CDS
LOCATION: 3..4334
US-08-670-707A-38

Query Match 3.9%; Score 46.6; DB 3; Length 4334;
Best Local Similarity 45.9%; Pred. No. 0.0059;
Matches 197; Conservative 0; Mismatches 229; Indels 3; Gaps 1;
QY 22 TTGTATCTCTCCCAAGTGGCGCTATGCCCTGGCATGCGACCGGACCATCCACAGT 81
DB 3840 TGTGATTAAACAGTTGCAGCATGCCCTGGGAATGCAGATAAAGCGATATCAGACTCA 3899
QY 82 GACATCTCTCTCCACCCGCTGGTGAACGCTCTGGACCCCGCTGTTCCACGCGCTACTGAGG 141

Query Match	3.1%	Score 37.6	DB 3	Length 432	
Best Local Similarity	44.1%	Pred. No. 0.52			
Matches 157	Conservative	0	Mismatches 199	Indels 0	Gaps 0
QY 119	CGCGCCGCCACAGCAGGTTGAGAGCAGTGACGGGGATGGGCGTGGTCCCCCGCAGGGT	178			
Db 9	CGGCCACGAGTCCGAAAGGGCGCGCATGGCCCCAGCAGGCCCGGAAAGGTGCGCGCTGGACC	68			
QY 179	CGGTGTTTCCCAAGGAGGAGGAGTACTTGCAGGTGGATCTACAACTGCTGCACCTGGTG	238			
Db 69	GGTGTGTCACCGGTGCTCGGGACCTGCAGGCCCGCCAGGTGCCTGGACCGGCTGG	128			
QY 239	CTCTGGTGGCACCCAGGACGGGATGCCGGGGGCGCTGGCCAGGAGGAGTCTCCCGAGGT	298			
Db 129	TCCACGGGTGCTCGGGACCTGCAGGCCCGCCAGGTGCCTGGACCGCGCTGGTCCACC	188			
QY 299	ACCGGCTGCGTTTACTCCCGGGATGGTGC CGCTGGATGGGCTGGAAGAGCCGCTGGGGTC	358			

```

RESULT 13
US-08-642-255-50
; Sequence 50, Application US/08642255
; Patent No. 5773249
; GENERAL INFORMATION:
; APPLICANT: CAPPELLO, Joseph
; APPLICANT: FERRARI, Franco A.
; TITLE OF INVENTION: High Molecular Weight Collagen-Like
; TITLE OF INVENTION: Protein Polymers
; NUMBER OF SEQUENCES: 135
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/642,255
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: ROWLAND, Bertram I.
; REGISTRATION NUMBER: 20,015
; REFERENCE/DOCKET NUMBER: A55556-3/BTR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 494-8700
; TELEFAX: (415) 494-8771
; TELEX: 910 277299 FHT UR
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 756 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "synthetic"
US-08-642-255-50

```

```

QY 299 ACCGGCTGCTTACTCCCGGATGTTGCGCGCTGGATGGCTGGAAGGACCGCTGGGGTC 358
Db 189 GGGTCTCCGGGACCTGACAGCGCCGACAGTGGCTGCGCTGACCGCTGGTCCACCGGGTC 248
QY 359 AGGAGTGATCTCAGGCAATGAGGACCTCAGGAGTGGTGGCTGAAGGACCTTGGGCCCC 418
Db 249 TCCGGACCTGACAGCGCCGACGCTGGCTGACCGCTGGTCCACCGGGTCTCCGGG 308
QY 419 CCATGGTCCCGACCTGCTTCTTACCCCGGGCTGACCGGGCTCATGAGCGTC 474
Db 309 ACCTGACGCGCCGACGCTGCGCTGACCGGCTGGTCCACCGGGTCTCCGGGAC 364

RESULT 14
US-08-672-569-1
; Sequence 1, Application US/08672569
; Patent No. 5827687
; GENERAL INFORMATION:
; APPLICANT: KOYAMA, No. 5827687uto
; APPLICANT: MIYOSHI, Ei-ji
; APPLICANT: IHARA, Yoshito
; APPLICANT: NISHIKAWA, Atsushi
; APPLICANT: TANIGUCHI, Naoyuki
; TITLE OF INVENTION: No. 5827687el Promoter and Method of Gene Expression
; TITLE OF INVENTION: Using the Same
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/672,569
; FILING DATE: 28 JUNE 1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: WEINER, Marc S.
; REGISTRATION NUMBER: 32,181
; REFERENCE/DOCKET NUMBER: 1422-0261P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1116 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-672-569-1

Query Match 3.1%; Score 37.2; DB 3; Length 1116;
Best Local Similarity 49.0%; Pred. No. 0.85;
Matches 99; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 201 GTACTTCAGGTGATCTACAAGCACTGCACCTGCTGGTGGCTGAGGACCGACCGGAGC 260
Db 709 GTACTGCGAGTGTAGCTTTTCCCAAGTCGTCTACTCTCTGGAGCAAGCTTAGCAGT 768
QY 261 GCATGCGGGGGCTGGGCAAGAGTTCCTCCCGGAGCTACCGGCTGCTTACTTCCCGGA 320
Db 769 GCAGCTTCACAGTCAAGTGGGGTGGTCCAGCGGAGAGCAAGCTCAGAGGGGGCAGGG 828

```

```

QY 321 TGGTCCGCGCTGGATGGCTGGAAGGACCGCTGGGGTCAAGGAGTGTATCTCAGGCAATGA 380
Db 829 TGGTGGCTCGGGATCTCAGGGAAGGGCTATGGAGACACGGCGTGTCTTCACTGCTGGG 888
QY 381 GGACCCTGAGGAGTGGTGTG 402
Db 889 GCTTTTCAGGGGCTTGGTACCG 910

RESULT 15
US-08-672-569-2
; Sequence 2, Application US/08672569
; Patent No. 5827687
; GENERAL INFORMATION:
; APPLICANT: KOYAMA, No. 5827687uto
; APPLICANT: MIYOSHI, Ei-ji
; APPLICANT: IHARA, Yoshito
; APPLICANT: NISHIKAWA, Atsushi
; APPLICANT: TANIGUCHI, Naoyuki
; TITLE OF INVENTION: No. 5827687el Promoter and Method of Gene Expression
; TITLE OF INVENTION: Using the Same
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/672,569
; FILING DATE: 28 JUNE 1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: WEINER, Marc S.
; REGISTRATION NUMBER: 32,181
; REFERENCE/DOCKET NUMBER: 1422-0261P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 817 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-672-569-2

Query Match 3.1%; Score 37.2; DB 3; Length 817;
Best Local Similarity 49.0%; Pred. No. 0.78;
Matches 99; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 201 GTACTTCAGGTGATCTACAAGCACTGCACCTGCTGGTGGCTGAGGACCGACCGGAGC 260
Db 410 GTACTGCGAGTGTAGCTTTTCCCAAGTCGTCTACTCTCTGGAGCAAGCTTAGCAGT 469
QY 261 GCATGCGGGGGCTGGGCAAGAGTTCCTCCCGGAGCTACCGGCTGCTTACTTCCCGGA 320
Db 470 GCAGCTTCACAGTCAAGTGGGGTGGTCCAGCGGAGAGCAAGCTCAGAGGGGGCAGGG 529
QY 321 TGGTCCGCGCTGGATGGCTGGAAGGACCGCTGGGGTCAAGGAGTGTATCTCAGGCAATGA 380

```


Db 530 TGGTGGCCTGGGATCTCAGGGAAGGGCTATGGGAGCACGGCGGTCTCTCAGTGCTGGG 589

Qy 381 GGACCCCTGAGGAGTGTGCTG 402

Db 590 GCTTTCAGGGGCTTGGTACCG 611

Search completed: November 4, 1999, 07:29:41
Job time: 6329 sec


```

Db 675 CGGCTCGTACTCCCGGATGGTCCCGCTGGATGGGCTGAAGAGCCGCTGGGGTCTAG 734
QY 361 GAGGTGATCTCAGGCAATGAGGACCTTGAAGAGTGGTCTGAAGAGACCTTTGGGCCCCC 420
Db 735 GAGGTGATCTCAGGCAATGAGGACCTTGAAGAGTGGTCTGAAGAGACCTTTGGGCCCCC 794
QY 421 ATGGTGGCCGAGTGGTTCCTTACCCCGGCTGACCGGGTCTAGGAGTCTGTCTG 480
Db 795 ATGGTGGCCGAGTGGTTCCTTACCCCGGCTGACCGGGTCTAGGAGTCTGTCTG 854
QY 481 CGGTAGAGCTCTATGGTTCCTTGAAGGATGGAGTCTGTCTTACACCGCCCTGTG 540
Db 855 CGGTAGAGCTCTATGGTTCCTTGAAGGATGGAGTCTGTCTTACACCGCCCTGTG 914
QY 541 GGGCAGACAATGATTTATCTAGGCGCTGTACCTCAACGACTCCACCTATGACGGACAT 600
Db 915 GGGCAGACAATGATTTATCTAGGCGCTGTACCTCAACGACTCCACCTATGACGGACAT 974
QY 601 ACCGTGGGCGGAGTGCAGTATGGGGTCTGGCCAGCTGGCAGATGGTGGTGGGGCTG 560
Db 975 ACCGTGGGCGGAGTGCAGTATGGGGTCTGGCCAGCTGGCAGATGGTGGTGGGGCTG 1034
QY 661 GATGACTTTAGGAAGAGTCAGAGTGGGGTCTGGCCAGCTATGACTATGTGGGATGG 720
Db 1035 GATGACTTTAGGAAGAGTCAGAGTGGGGTCTGGCCAGCTATGACTATGTGGGATGG 1094
QY 721 ACACACACAGCTTCTCAGTGGCTATGTGAGATGGAGTGGAGTGGAGTGGAGTGGAG 780
Db 1095 ACACACACAGCTTCTCAGTGGCTATGTGAGATGGAGTGGAGTGGAGTGGAGTGGAG 1154
QY 781 GCCTCCAGGCTATCAGGTCACCTGTACACATGACACAGCTGGGAGCCGCTGTGCT 840
Db 1155 GCCTCCAGGCTATCAGGTCACCTGTACACATGACACAGCTGGGAGCCGCTGTGCT 1214
QY 841 GCGGGGTGGAATGTCCTTCGCGGCTGGCCCTTGCATGGCTGGGAGGGGAGCCCATG 900
Db 1215 GCGGGGTGGAATGTCCTTCGCGGCTGGCCCTTGCATGGCTGGGAGGGGAGCCCATG 1274
QY 901 GCGCACACCTAGGGGGCAACCTGGGGGAGCCAGAGCCCGGGGTGTCTCAGTCCGCCCT 960
Db 1275 GCGCACACCTAGGGGGCAACCTGGGGGAGCCAGAGCCCGGGGTGTCTCAGTCCGCCCT 1334
QY 961 GCGCGCTGTGGCTTCCTTCAGTGGCCGCTTCCCTTCTTGGGGGCTGTGTTACTC 1020
Db 1335 GCGCGCTGTGGCTTCCTTCAGTGGCCGCTTCCCTTCTTGGGGGCTGTGTTACTC 1394
QY 1021 TTCAGCAAAATCTCCTTCATCTCTGATGTGGTGAACAAATTCCTTCGCGCACTGGGAGG 1080
Db 1395 TTCAGCAAAATCTCCTTCATCTCTGATGTGGTGAACAAATTCCTTCGCGCACTGGGAGG 1454
QY 1081 ACCTTCCCGCAGCCCTGTGGCCGCTGGCCGCTGGCCGCTGGCCGCTGGCCGCTGG 1140
Db 1455 ACCTTCCCGCAGCCCTGTGGCCGCTGGCCGCTGGCCGCTGGCCGCTGGCCGCTGG 1514
QY 1141 GAGCTGGAGCCAGAGGCGCAGCCGCTGGCCGCTGGCCGCTGGCCGCTGGCCGCTGG 1197
Db 1515 GAGCTGGAGCCAGAGGCGCAGCCGCTGGCCGCTGGCCGCTGGCCGCTGGCCGCTGG 1571

```

RESULT 3

```

T93785
ID T93785 standard; cDNA; 3962 bp.
AC T93785.
DT 16-FEB-1998 (first entry)
DE Human mammary carcinoma kinase 10 (MCK-10) cDNA sequence.
KW Mammary carcinoma kinase; MCK-10; receptor tyrosine kinase;
KW proliferative disease; cancer; insulin receptor family;
KW tyrosine kinase neurotrophin receptor; MCK-10 activity;
KW neurological disorder; aberrant expression; ds.
OS Homo sapiens.
FH Key
FT Location/Qualifiers
321..3077

```

FT PN US5677144-A. /*tag= a

14-OCT-1997. 336343.
 08-NOV-1994; 336343.
 16-NOV-1993; US-153397.
 (ALVE/) ALVES F H E.
 (ULLR/) ULLRICH A.
 PI ALVES FHE, ULLRICH A.
 DR WPI; 97-511869/47.
 DR P-PSDB; W34672.
 PT Truncated receptor tyrosine kinase CCK-2 - and nucleic acid coding
 for it, useful for cancer diagnosis
 PT Disclosure; Fig 1; 70pp; English.
 PS The present sequence represents the cDNA of a mammary carcinoma kinase,
 CC called MCK-10. This kinase belongs to a novel family of receptor tyrosine
 CC kinases, and expression is associated with proliferative diseases such as
 CC cancer. The MCK-10 receptor tyrosine kinase has extensive sequence
 CC similarity to the insulin receptor family. The MCK-10 gene was isolated
 CC by PCR using 2 degenerate oligonucleotide primer pools, using a template
 CC cDNA synthesised by reverse transcription of poly-A RNA from the human
 CC mammary carcinoma cell line MCF7. MCK-10 is expressed in brain tissue,
 CC and the protein shares homology with the tyrosine kinase neurotrophin
 CC receptor. Modulation of MCK-10 activity therefore may be used for
 CC treatment of neurological disorders. MCK-10 is also expressed in a
 CC variety of cancer cell lines and tumour tissue. The present sequence,
 CC or parts of it, can be used for diagnostic purposes to detect aberrant
 CC expression of MCK-10 genes. Inhibitors of MCK-10 receptor activity may
 CC have therapeutic value in the treatment of diseases such as cancer.
 SQ Sequence 3962 BP; 735 A; 1234 C; 1182 G; 811 T;

Query Match

Best Local Similarity 99.6%; Score 1192.2; DB 1; Length 3962;
 Matches 1194; Conservative 99.7%; Pred. No. 6.4e-274;
 Mismatches 3; Indels 0; Gaps 0;

```

QY 1 GATGCTGACATGAGGAGACATTTTATCTGCTGCAAGTGCCTGATGCCCTGGCATGTCAG 60
Db 375 GATGCTGACATGAGGAGACATTTTATCTGCTGCAAGTGCCTGATGCCCTGGCATGTCAG 434
QY 61 GACCGGACCATCCAGACAGTGCATCTCTGCTTCCAGTCTCTGCTGCTGCTGCTGCTGCTG 120
Db 435 GACCGGACCATCCAGACAGTGCATCTCTGCTTCCAGTCTCTGCTGCTGCTGCTGCTGCTG 494
QY 121 GCGCGCCACAGCAGGTGGAGAGCAGTGCAGGGGATGGGGCTGGTGGCCCGCAGAGGTGCG 180
Db 495 GCGCGCCACAGCAGGTGGAGAGCAGTGCAGGGGATGGGGCTGGTGGCCCGCAGAGGTGCG 554
QY 181 GTGTTTCCCAAGGAGGAGGAGTACTTGCAGTGGATGTACACAGACTGCACCTGGTGGCT 240
Db 555 GTGTTTCCCAAGGAGGAGGAGTACTTGCAGTGGATGTACACAGACTGCACCTGGTGGCT 614
QY 241 CTGGTGGGACCCAGGAGCGCATGCCGGGGCTGGGCAAGGAGTTCCTCCCGGAGCTAC 300
Db 615 CTGGTGGGACCCAGGAGCGCATGCCGGGGCTGGGCAAGGAGTTCCTCCCGGAGCTAC 674
QY 301 CGGCTCGCTTACTCCCGGATGGTCCGCTGGATGGGCTGGAAGGAGCCGCTGGGGTCTAG 360
Db 675 CGGCTCGCTTACTCCCGGATGGTCCGCTGGATGGGCTGGAAGGAGCCGCTGGGGTCTAG 734
QY 361 GAGGTGATCTCAGGCAATGAGGACCTGAGGAGTGGTGTGTAAGGACCTTTGGGCCCCC 420
Db 735 GAGGTGATCTCAGGCAATGAGGACCTGAGGAGTGGTGTGTAAGGACCTTTGGGCCCCC 794
QY 421 ATGGTTGGCCGAGTGGTTCGCTTCTACCCCGGCTGACCGGGTCTATGAGCTCTGTCTG 480
Db 795 ATGGTTGGCCGAGTGGTTCGCTTCTACCCCGGCTGACCGGGTCTATGAGCTCTGTCTG 854
QY 481 CGGGTAGAGCTCTATGGCTGGCTCTGGAGGATGGAGTCTGTCTTACACCGCCCTGTG 540
Db 855 CGGGTAGAGCTCTATGGCTGGCTCTGGAGGATGGAGTCTGTCTTACACCGCCCTGTG 914
QY 541 GGGCAGACAATGATTTATCTGAGGCGCTGTACCTCAACGACTCCACCTATGACGGACAT 600

```

Db 915 GGCAGACAAATGATTTATCTGAGCGCGTGTACTCTCAAGACTCCACCTATGACGACAT 974
Qy 601 ACCGTGGCGGACTGCAGTATGGGGTCTGGCCAGCTGACAGTGTGTGGGGCTG 660
Db 975 ACCGTGGCGGACTGCAGTATGGGGTCTGGCCAGCTGACAGTGTGTGGGGCTG 1034
Qy 661 GATGACTTTAGGAAGAGTCAAGAGCTGCGGGTCTGGCCAGCTATGACTATGTGGATG 720
Db 1035 GATGACTTTAGGAAGAGTCAAGAGCTGCGGGTCTGGCCAGCTATGACTATGTGGATG 1094
Qy 721 AGCAACACAGCTTCTCCAGTGGCTATGTGGAGATGAGTTTGTAGTTTACCGGCTGAG 780
Db 1095 AGCAACACAGCTTCTCCAGTGGCTATGTGGAGATGAGTTTGTAGTTTACCGGCTGAG 1154
Qy 781 GCCTTCCAGGCTATGACAGTCCACTGTAAACAATGACACAGCTGGAGCGGCTGTGCT 840
Db 1155 GCCTTCCAGGCTATGACAGTCCACTGTAAACAATGACACAGCTGGAGCGGCTGTGCT 1214
Qy 841 GCGGGGTGGAATGCTCTCCGCGTGGCCCTGCCATGCGCTGGAGGGGAGCCCATG 900
Db 1215 GCGGGGTGGAATGCTCTCCGCGTGGCCCTGCCATGCGCTGGAGGGGAGCCCATG 1274
Qy 901 CGCACAACTAGGGGCAACCTGGGGAGCCAGAGCCCGGCTGTCTCAGTCCCGCTT 960
Db 1275 CGCACAACTAGGGGCAACCTGGGGAGCCAGAGCCCGGCTGTCTCAGTCCCGCTT 1334
Qy 961 GCGGCGGTGGCTGCTCTTCTGACGTGGCTTCTCTTTGGGGGCGCTGTACTC 1020
Db 1335 GCGGCGGTGGCTGCTCTTCTGACGTGGCTTCTCTTTGGGGGCGCTGTACTC 1394
Qy 1021 TTCAGCAAACTCTCTTCTATCTGATGTGGTGAACAATTCCTCTCGGCACTGGAGGC 1080
Db 1395 TTCAGCAAACTCTCTTCTATCTGATGTGGTGAACAATTCCTCTCGGCACTGGAGGC 1454
Qy 1081 ACCTTCGCGGCGGCTGTGGTGGCGGCTGGCCAGCTTCCACCACTTCACAGCTTG 1140
Db 1455 ACCTTCGCGGCGGCTGTGGTGGCGGCTGGCCAGCTTCCACCACTTCACAGCTTG 1514
Qy 1141 GAGTGGAGCCAGAGCGGCGTGGCCAGCGGCGGAGGGAGCCCGGCGG 1197
Db 1515 GAGTGGAGCCAGAGCGGCGTGGCCAGCGGCGGAGGGAGGGAGCCCGGCGG 1571

RESULT 4

ID Q84782
AC Q84782 standard; DNA; 3754 BP.
DE 17-AUG-1995 (first entry)
DT Protein-tyrosine-kinase PTK22.
KW Protein-tyrosine-kinase; PTK; disoidin domain receptor; cancer;
KW breast tumor; mammary carcinoma; diagnosis; prognosis; therapy; ds.
OS Homo sapiens.
FH Key Location/Qualifiers
FT cds 142..2886
FT /*tag= a
FN W09502187-A.
PD 19-JAN-1995.
PF 08-JUL-1994; G01480.
PR 09-JUL-1993; GB-014271.
PA (CANC-) CANCER RES INST.
PA (WELL) WELLCOME FOUND LTD.
PI Barker KT, Crompton MR, Gusterson BA, Martindale JE;
PI Mitchell PJ, Page M, Spence P;
DR WPI: 95-066991/09.
DR P-PSDB: R71100.
PT Method for screening substances, using protein tyrosine kinase -
PT for potential utility as therapeutic agents for cancer
PS Claim 1; Page 26-30; 51pp; English.
CC cDNA derived from tumor metastatic tissue was amplified using
CC primers (given in Q84783-84) based on sequences (R71101, R71103)
CC associated with protein-tyrosine-kinases (PTK). Novel PTK22 was
CC identified in an isolated subclone. The 3' sequence of PTK22 was
CC obtained by reverse transcription (using the primer of Q84786) and

CC PCR amplification (primers Q84787-88) of RNA of human breast
CC carcinoma cell line MDA MG 468. The partial DNA sequence of PTK22
CC is given in Q84782.
SQ Sequence 3754 BP; 713 A; 1145 C; 1121 G; 775 T;
Query Match 97.8%; Score 1171.2; DB 1; Length 3754;
Best Local Similarity 99.1%; Pred. No. 6e-269;
Matches 1189; Conservative 0; Mismatches 8; Indels 3; Gaps 1;
Qy 1 GATGCTGACATGAGGGACATTTTATCTCCCTGCTCCCAAGTGGCCCTATGCTGGGATCGAG 60
Db 196 GATGCTGACATGAGGGACATTTTATCTCCCTGCTCCCAAGTGGCCCTATGCTGGGATCGAG 255
Qy 61 GACCGGACATCCAGACAGTGCATCTCTCTCCAGCTCCTGCTCAGATTCACCTGCC 120
Db 256 GACCGGACATCCAGACAGTGCATCTCTCTCCAGCTCCTGCTCAGATTCACCTGCC 315
Qy 121 GCGCGCCACACAGAGTTGGAGAGAGTGCAGCGGATGGGCTGCTGGCCCGAGGGTCG 180
Db 316 GCGCGCCACACAGAGTTGGAGAGAGTGCAGCGGATGGGCTGCTGGCCCGAGGGTCG 375
Qy 181 GTGTTTCCCAAGGAGGAGTACTTCCAGTGGATCTACAACGACTGCACCTGGTGCCT 240
Db 376 GTGTTTCCCAAGGAGGAGTACTTCCAGTGGATCTACAACGACTGCACCTGGTGCCT 435
Qy 241 CTGTTGGGACCCAGGAGCGATGCCGGGCTGGGCAAGAGTCTTCCCGGAGGTAC 300
Db 436 CTGTTGGGACCCAGGAGCGATGCCGGGCTGGGCAAGAGTCTTCCCGGAGGTAC 495
Qy 301 CGGCTGCGTTACTCCCGGATGGTGGCGCTGGATGGGCTGGAAGACCGCTGGGGTCAG 360
Db 496 CGGCTGCGTTACTCCCGGATGGTGGCGCTGGATGGGCTGGAAGACCGCTGGGGTCAG 555
Qy 361 GAGTGTGATCTCAGCAATGAGGACCTGAGGAGTGGTCTGAAGGACCTTGGGCCCCC 420
Db 556 GAGTGTGATCTCAGCAATGAGGACCTGAGGAGTGGTCTGAAGGACCTTGGGCCCCC 615
Qy 421 ATGTTGGCCGACTGTTGCTTCTTACCCCGGCTGACCGGGTCTAGGAGCTCTGTCTG 480
Db 616 ATGTTGGCCGACTGTTGCTTCTTACCCCGGCTGACCGGGTCTAGGAGCTCTGTCTG 675
Qy 481 CGGCTAGAGCTCTATGGCTGCTCTGGAGGATGAGTCTCTTACACCGCCCTGTG 540
Db 676 CGGCTAGAGCTCTATGGCTGCTCTGGAGGATGAGTCTCTTACACCGCCCTGTG 735
Qy 541 GGCAGACAAATGATTTATCTGAGGCGCTGTACCTCAACGACTCCACCTATGAGGACAT 600
Db 736 GGCAGACAAATGATTTATCTGAGGCGCTGTACCTCAACGACTCCACCTATGAGGACAT 795
Qy 601 ACCGTGGCGGACTGCAGTATGGGGTCTGGGCCAGCTGGCAGATGTTGTTGGGGCTG 660
Db 796 ACCGTGGCGGACTGCAGTATGGGGTCTGGGCCAGCTGGCAGATGTTGTTGGGGCTG 855
Qy 661 GATGACTTTAGGAAGAGTCAAGAGCTGCGGGTCTGGCCAGGCTATGACTATGTGGATG 720
Db 856 GATGACTTTAGGAAGAGTCAAGAGCTGCGGGTCTGGCCAGGCTATGACTATGTGGATG 915
Qy 721 AGCAACACAGCTTCTCCAGTGGCTATGTGGAGATGAGTTTGTAGTTTACCGGCTGAG 780
Db 916 AGCAACACAGCTTCTCCAGTGGCTATGTGGAGATGAGTTTGTAGTTTACCGGCTGAG 975
Qy 781 GCCTTCCAGGCTATGACAGTCCACTGTAAACAATGACACAGCTGGAGCGGCTGTGCT 840
Db 976 GCCTTCCAGGCTATGACAGTCCACTGTAAACAATGACACAGCTGGAGCGGCTGTGCT 1035
Qy 841 GCGGGGTGGAATGCTCTCCGCGTGGCCCTGCCATGCGCTGGAGGGGAGCCCATG 900
Db 1036 GCGGGGTGGAATGCTCTCCGCGTGGCCCTGCCATGCGCTGGAGGGGAGCCCATG 1095
Qy 901 CGCACAACTAGGGGCAACCTGGGGAGCCAGAGCCCGGCTGTCTCAGTCCCGCTT 960
Db 1096 CGCACAACTAGGGGCAACCTGGGGAGCCAGAGCCCGGCTGTCTCAGTCCCGCTT 1155

```

QY 961 GCGGCGCGTGGCTCGCTTTCTGAGTGGCGCTTCTTTTGGGGGCGCTGTTACTC 1020
Db 1156 GCGGCGCGTGGCTCGCTTTCTGAGTGGCGCTTCTTTTGGGGGCGCTGTTACTC 1215
QY 1021 TTCACGGAATCTCTTTCATCTCTGATGTGTGAGCAATTCCTCTCCGGCACTGGAGGC 1080
Db 1216 TTCACGGAATCTCTTTCATCTCTGATGTGTGAGCAATTCCTCTCCGGCACTGGAGGC 1275
QY 1081 ACCTTCCCGCCAGCCCTTGGTGGCGCTGGCCACCTCCACCACTTCAGCAGCTTG 1140
Db 1276 ACCTTCCCGCCAGCCCTTGGTGGCGCTGGCCACCTTCACCACTTCAGCAGCTTG 1335
QY 1141 GAGCTGGAGCCAGAGCCAGCAGC---AGCCGTGGCCAAAGCCCGAGGGAGCCCGACCGCC 1197
Db 1336 GAGCTGGAGCCAGAGCCAGCAGCAGCAGCCGCTGGCCAAAGCCCGAGGGAGCCCGACCGCC 1395

RESULT 5
Q92521
ID Q92521 standard; cDNA to mRNA; 3157 BP.
AC Q92521;
DT 26-NOV-1995 (first entry)
DE Human colonic adenocarcinoma kinase 2 (CKK-2) cDNA.
KW Mammary carcinoma kinase 10; MCK-10; transmembrane receptor; CKK-2;
OS Homo sapiens.
FH Key
FT cds
FT Location/Qualifiers
   370..2937
   /*tag= a
PN W09514088-A.
PD 26-MAY-1995.
PE 16-NOV-1994; E03797.
PR 16-NOV-1993; US-153397.
PA (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
PI Alves PHE, Ullrich A;
DR WPI: 95-224054/29.
DR P-PSDB; R75503.
PT New nucleic acid encoding MCK-10 receptor tyrosine kinase - and
PT derived vectors, transformed cells, proteins and antibodies useful
PT for diagnosis and treatment of proliferative disease, esp. cancer,
PT and for screening modulators
PS Disclosure: Page 56-57; 115pp; English.
CC A member of the mammary carcinoma kinase 10 (MCK-10) receptor
CC tyrosine kinase family was identified using a PCR (with two degenerate
CC oligo primer pools based on conserved sequences of the kinase domain of
CC receptor tyrosine kinases) and cDNA prep. from colonic
CC adenocarcinoma RNA. The nt sequence of the novel receptor,
CC designated CKK-2, is given in Q92521 and the deduced AA sequence in
CC R75503. Analysis of the CKK-2 nt and AA sequence indicated
CC significant homology with MCK-10 throughout the extracellular,
CC transmembrane and intracellular regions. The regions of homology
CC extend into the N-terminus consensus sequence for the discoidin I
CC like family of proteins.
SQ Sequence 3157 BP; 790 A; 802 C; 759 G; 806 T;

Query Match 27.4%; Score 327.4; DB 1; Length 3157;
Best Local Similarity 58.8%; Pred. No. 6.7e-69;
Matches 644; Conservative 0; Mismatches 436; Indels 15; Gaps 4;

QY 13 AAGGGACATTTTGCATGCTCCAGTGCCTGATGCGCTGGGATGCGAGCCGACCATC 72
Db 433 AAGGCTCAGGTTAATCCAGCTATATCCGCTATCTCTGGCATGTCAGGAGGCCAGATT 492
QY 73 CCAGACAGTGCATCTCTGCTCCAGCTCTGTCAGATTCCTCCGCGCCGCCACAGC 132
Db 493 CCAGATGAGGACATCAGCATCTCCAGCTCAGTCAGATGTCAGAGCTCCCAATATGGA 552
QY 133 AGGTGGAGAGCAGTGAAGGGGATGGGCTGGTGGCCGAGGCTGGTGGTGGTGGTGG 192
Db 553 AGGCTGGAAGTGAAGAGGGGATGAGGCTGGTGGCCCTGAGATTCAGTGAACCTGAT 612

```

```

QY 193 GA---GGAGGAGTACTTGCAGGTGGATCTACAACGACTCCACCTGGTGGCTGCTGGGC 249
Db 613 GACCTGAAGGAGTTTCTGCAGATTGACATGGCACCCCTCCATTTTATCATTCTGTGGGG 672
QY 250 ACCAGGGAGCGATCCCGGGGCGCTGGGCAAGAGGTTCTCCCGGAGTACCGGCTGGT 309
Db 673 ACCAGGGGCGCATGCGAGAGGTGATGCAATCGAGTTGGCCCCCATGTACAAATCAAT 732
QY 310 TACTCCCGGATGGTGGCGCTGGATGGGCTGGAAGGACCGCTGGGGTTCAGAGGTGATC 369
Db 733 TACAGTCGGGATGGCACTCGCTGGATCTCTGGCGGAACCGTCATGGGAACAGGTGCTG 792
QY 370 TCAGGCAATGAGGACCGCTGAGGAGTGGTCTGAAGGACCTTGGGCCCCCCTGTTGCC 429
Db 793 GATGGAATTAAGAACCCCTATGACATTTTCTTAAAGGACTTGGAGCGCCCATTTGAGCC 852
QY 430 CGACTGGTTCGCTTTCACCCCGGCTGACCGGGTCAATGAGCGTCTGTCTGGGGTAGAG 489
Db 853 AGATTTGTCGGGTTCAATTCAGTCAACCGACCACTCCATGAATGTGTATGAGAGTGGAG 912
QY 490 CTCTATGGTCTCTGAGGAGTGAATCTCTGTCTTACACCGCCCTGTGGGGAGACA 549
Db 913 CTTTACGGCTGTCTGCTGGCTAGATGGCTTGGTGTCTTACAATGCTCCAGCTGGCAGCAG 972
QY 550 ATGATATTATCTGAGC-----CGTGTACCTCAAGGACTCCACCTATCAGGACATACC 603
Db 973 TTTGTACTCCCTGGAGGTTTCCATATTATCTGAATGATTTCTCTATGATGAGCTGTT 1032
QY 604 GTGGCGGACTGCAGTATGGGGTCTGGGCGACGCTGGAGATGGTGTGGTGGGGTGGAT 663
Db 1033 GGATACAGTATGACAGAGGG---CTAGGCCAATTGACCGATGGTGTCTGGCCTGGAC 1089
QY 664 GACTTTAGGAAGAGTCAAGGAGTGGCGGTCTGGCCAGGCTATGACTATGTGGAGGAGC 723
Db 1090 GATTTCCACCCAGACCCCATGAATACCACGCTGTGGCCCGGCTATGACTATGTGGCTGCGG 1149
QY 724 AACACAGCTTCTCCAGTGGCTATGAGATGGAGTGGTTTGTACCGGCTGAGGGCC 783
Db 1150 AAGGAGAGTGGCCCAATGGCTACATTTGAGATCATGTTTGAATTTGACCGCATCAGGAAT 1209
QY 784 TTCAGGCTATGCAAGTCCACTGTAACACATGCACACGCTGGGAGCCGCTCTGCTGCTGGC 843
Db 1210 TTCACTACCATGAGGTCCACTGCACACACATGTTTGTCTAAAGTGTGAGATCTTTAAG 1269
QY 844 GGGGTGAATGTCCGTTCCGGCTGGCCCTGCCATGCCCTGGAGGGGAGCCCATGCGC 903
Db 1270 GAGGTACAGTCTACTTCCGCTCTGAA--GCCAGTGAAGTGGGAACCTAATGCCATTTC 1326
QY 904 CACNACCTAGGGGCAACCTGGGGGACCCAGAGCCCGGCTGTCTCAGTGGCCCTTGGC 963
Db 1327 TTCCCGCTTCTCTGGATGACGTCACCCCGAGTGTCTGGTTCACGGTGGCTCTCCAC 1386
QY 964 GGCGGTGTGCTGCTGCTTTCGAGTGGCGCTTCTCTTTCGGGGGCGCTGTTACTTCTTC 1023
Db 1387 CACCGAATGGCCAGTGCATCAAGTGTCAATACCATTTTGCAGATACCTGGATGATGTT 1446
QY 1024 AGCGAATCTCTTCAATCTGATGTGGTGAACAATTCCTCTCCGGCACTGGGAGCAC 1083
Db 1447 AGTGAGATCACTTCCCAATCAGATGCTGCAATGTACAACTCTGAAGCCCTGCCACC 1506
QY 1084 TTCCCGCCAGCCGCC 1098
Db 1507 TCTCTATGGCACCC 1521

```

RESULT 6

Q92523
ID Q92523 standard; cDNA to mRNA; 3157 BP.

AC Q92523;
DT 26-NOV-1995 (first entry)
DE Human colonic adenocarcinoma kinase 2 (CKK-2) cDNA.
KW Mammary carcinoma kinase 10; MCK-10; transmembrane receptor; CKK-2;
KW receptor tyrosine kinase; colonic adenocarcinoma kinase 2; ss.

CC aberrant expression of CCK-2 genes. Engineered cell lines, containing
CC recombinant vectors with the present sequence, are useful for producing
CC infectious retroviral particles. The cell lines may also be used to
CC evaluate and screen drugs involved in CCK-2 activation and regulation.
SQ Sequence 317 BP; 790 A; 802 C; 759 G; 806 T;

Query Match	27.4%	Score 327.4;	DB 1;	Length 3157;
Best Local Similarity	58.8%;	Pred. No. 6.7e-69;		
Matches 644;	Conservative 0;	Mismatches 436;	Indels 15;	Gaps 4;

QY	13	AAGGGACATTTTGATCTCGCAAGTCGCGCTATGCCCTGGCATGCAGNACCGGACCATT	C 72
DB	433	AAAGCTCAGGTTAATCCAGCTATATGCCGCTATCTCTGGGCATGTTCAGAGGCCAGATT	492
QY	73	CCAGCAGTGCATCTCTGCTTCCAGTCTCTGTGCAGATTCACACTGCGCGCCACACAG	C 132
DB	493	CCAGATGAGGACATCACAGCTTCCAGTCAGTGGTTCAGAGTCCACAGTGCACAAATATGA	552
QY	133	AGTGTGGAGCAGTGCAGGGGATGGGCGCTGGTGGCCCCAGGTCGGTGTTCCTCCAG	C 192
DB	553	AGCTGGACTCAGAAGAGGGATGGAGCTGGTGGCCCTGAGATTCCAGTGGAACTGAT	612
QY	193	GA---GGAGGACTACTTGCAGTGGATACAACGACTGCACCTGTGTGGCTCTGTGGGC	C 249
DB	613	GACCTGAAGGATTTCTGCAGATTGACTTGCACACCTCCATTTATCACTCTGTGGGG	672
QY	250	ACCCAGGAGCGATGCGCGGGGCTGGGCAAGAGTCTTCCGGAGCTACCGGTGCGT	C 309
DB	673	ACCCAGGGGCGCATGCAAGAGGTCTATGGCATGCAAGTTTCCGCCCAGTACAAAGATCAAT	732
QY	310	TACTCCCGGATGGTTCGCGCTGGATGGCTTGAAGACCGCTGGGTGCAGGAGTGATC	C 369
DB	733	TACAGTCGGATGGCACTCGCTGGATCTCTGGCGGAACCGTCAATGGGAACAGTGTCTG	792
QY	370	TCAGGCAATGAGGACCTTGAGGAGTGGTGTGAAGACCTTGGGCCCCCATGTTGCC	C 429
DB	793	GATGGAAATAGTAACCCCTATGACATTTCTTAAGGACTTGGAGCGGCCCATTTGAGCC	852
QY	430	CGACTGGTTTCGCTTCTACCCCGGCTGACCGGGTCAATGAGCGTCTGTGGGGTAGAG	C 489
DB	853	AGATTGTCGGTTCATTCACAGTCACCGACCACTCCATGATGTGTATGAGATGGAG	912
QY	490	CTCATAGGTCCTCTGAGGGATGGACTCTCTGTTCACACCGCCCTCTGGGGCTGGAT	C 549
DB	913	CTTTACGGCTGTCTGGCTAGATGGTGTGGTGTCTTACAACTGCTCCAGTGGCGCAG	972
QY	550	ATGTATTTATCTGAGGC-----CGTGTACTCAAGCACTCCACCTATCAGGCAATACC	C 603
DB	973	TTTGTACTCCCTGGAGGTTCCATCATTTATCTGAATGTTCTCTATATGAGCGCTGT	1032
QY	604	GTGGGGGACTGCAGTATGGGGTCTGGGCCAGCTGGCAGATGGTGTGGGGCTGGAT	C 663
DB	1033	GGATACAGCATGACAGAAGG--CTAGGCCAATTGACCACTGGTGTGTCTGSCCTGGAC	1089
QY	664	GACTTTAGAGAGTCAGGACCTCGGGTCTGGCCAGGCTATGACTATGTGGGATGGAGC	C 723
DB	1090	GATTTACCCAGACCATGAATACCACTGTGGCCGGGTATGACTATGTGGGCTGGGG	1149
QY	724	AACCAAGCTTCTCAGTGGCTATGTGGAGATGGAGTTTGAGTTTGACCGGCTGAGGCC	C 783
DB	1150	AACGAGATGGCACCATGGCTACATTGAGATCATGTTTGAATTTGACCGCATCAGGAT	1209
QY	784	TTCCAGGCTATGCAGGTCCACTGTAAACAACATGCACACGCTGGGAGCCGCTGCTGCC	C 843
DB	1210	TTTCACTACCATGAAGGTCACCTGCAACAACATGTTTGTCTAAAGGTGTGAAGATCTTTAAG	1269
QY	844	GGGGTGGATGTTCGCTTCCGCGTGGCCCTGCCATGGCTGGGAGGGGAGCCATGGC	C 903
DB	1270	GAGTACAGTGTCTACTCCGCTCTGA--GCCATGATGGGAACCTAATGCCATTTC	1346
QY	904	CACAACTAGGGGGCAACTGTGGGGACCCAGAGCCCGGCTGTCTCAGTGGCCCTTGGC	C 963

Db	1327	TTCCCCCTTGCTCTCGATGACGTCAACCCCAAGTCTCGGTTGTTCACGGTGCCCTCTCCAC	1388
Qy	964	GGCGTGTTGGCTCGCTTTCTTCAGTGGCGCTTCCTCTTTGGGGGGCCCTGGTTACTCTTC	1023
Db	1387	CACGAATGCCAGTCCCATCAAGTGTCAATACCATTTGACAGATACTGGATGATCTTC	1446
Qy	1024	AGGAAATCTCTTCATCTCTGATGGTGTGAACAATTCCTCTCGGCACATGGGAGGCACC	1083
Db	1447	AGTGAGATCACCTTCCAATTCAGATGCTCAATGTACAACAACCTCTGAAGCCCTGCCACC	1506
Qy	1084	TTCCCGCCAGCCCCC	1098
Db	1507	TCTCCTATGGCACCC	1521
RESULT	8		
V48292			
IC	AD	V48292 standard; cDNA; 3096 BP.	
AC	V48292;		
DT	16-NOV-1998	(first entry)	
DE	Discoidin domain receptor 2 gene.		
KW	Discoidin domain receptor; transformation; metastasis; collagen; ss;		
KS	Cleidocranial dysplasia; Sickler syndrome; extracellular matrix; MMP-1		
OS	Homo sapiens.		
PH	Key	Location/Qualifiers	
FT	CDS	354..2921	
FT		/*tag= a	
FT		/product= "Discoidin domain receptor"	
FT	sig_peptide	354..416	
FT		/*tag= b	
FT	mat_peptide	417..2918	
FT		/*tag= c	
PN	W09834954-A2.		
PD	13-AUG-1998.		
PF	05-FEB-1998; CA0093.		
PR	06-FEB-1997; US-041578.		
PA	(MOUN) MOUNT SINAI HOSPITAL CORP.		
PI	Pawson A, Vogel W;		
DR	WPI; 98-447168/38.		
DR	P-PSDB; W77114.		
DR	Novel ligands of discoidin domain receptor tyrosine kinase,		
PT	especially collagen - useful for treating e.g. metastasis,		
PT	cleidocranial dysplasia or Sickler syndrome		
PS	Disclosure; Fig 22a; 115pp; English.		
CC	The DRR can be used to identify and evaluate substances which affect DRR		
CC	receptor tyrosine kinase signalling pathways in the cell. Compounds		
CC	which modulate such signalling pathways can be used to alter		
CC	transformation or metastasis in mammals, to treat conditions involving		
CC	structural or functional deregulation of collagens, e.g. Cleidocranial		
CC	dysplasia or Sickler syndrome, conditions requiring modulation of		
CC	extracellular matrix synthesis, degradation or remodelling, or to treat		
CC	conditions needing modulation of MMP-1 expression such as wound healing		
CC	Sequence 3096 BP; 762 A; 791 C; 752 G; 791 T;		
SQ			

Query Match	27.4%	Score 327.4	DB 1	Length 3096
Best Local Similarity	58.8%	Pred. No. 6.7e-69		
Matches 644	Conservative 0	Mismatches 436	Indels 15	Gaps 4
QY 13	AAGGGACATTTTGATCTCGCCAAAGTCGCGGTATGCGCTTGGCGCATGCAGGACCGGACCATC	72		
Db	417	AAAGCTCAGGTTTAATCCACAGCTATATGCGGTATCTCTGGGCGATGTCAGGAGSCCAGATT	476	
QY 73	CCACAGAGTGACATCTCTGCTTCAGAGTCCCTGGTCAGATCTCACTGCCGCCGCCACAGC	132		
Db	477	CCAGATGAGGACATCACAGCTTCACGTAGTGGTCAGAGTCCACAGCTGCCAANAATGGA	536	
QY 133	AGGTTGGACAGACATGACGGGATGCGGGCTGTGGTGTGCCCCCAGCGGTGGTGTTCCTCCAAG	192		
Db	537	AGGCTGGACTCAGAGAAGGGGATGGAGSCTGTGTCCTGAGATTCACGTGGAAACCTGAT	596	
QY 193	GA---GGAGGAGTACTTTCACAGTGGATCTACACAGCTGCACCTGCTGGCTCTGTGTGGGC	249		

```

Db 597 GACCTGAAGAGTTTCTCAGATTGACTTGCACACCTCCATTTTATCATCTCGTGGG 656
QY 250 ACCCAGGACGCATCCGGGGCTGGCAAGAGTTCTCCGGAGCTACCGCTCGCT 309
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 657 ACCCAGGGCCCATCAGAGGTCATGCGATCGATTGCCCCCATGTACAAGTCAAT 716
QY 310 TACTCCCGGATGTCGCGCTGGATGGGTGGAGGACCGCTGGGTGAGAGGTGATC 369
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 717 TACACTCGGATGCGACTCGCTGGATCTCTTGGGGAAACCGCTCATGGGAAACAGGTGCTG 776
QY 370 TCAGCATAGGACCTGAGGAGTGTGCTGAAGGACCTTGGGCCCCCATGTTGCC 429
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 777 GATGAAATAGTAACCCCTATGACATTTTCTAAAGACTTGGAGCGCCCATGTAGCC 836
QY 430 CCACTGGTTCGCTTCTACCCCGGCTCACCGGTCATGAGCGTCTGCTCGGGTAGAG 489
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 837 AGATTTCGGGTTCATTCAGTACCACACATCCCATGATGTGTATGAGAGTGGAG 896
QY 490 CTCTATGCTGCTCTGAGGAGTGAATCTGCTTACACCGCCCTGTGGGCGAGACA 549
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 897 CTTTACGGCTGTGCTGCTAGTAGCTTGGTGTCTTAAATGCTCCAGCTGGGCGAG 956
QY 550 ATGTAATTTATCTAGGC-----CGTGTACCTCAACGACTCCACCTATGACGGACATACC 603
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 957 TTTGTACTCCCTGGAGGTTCCATCATTTATCTGAATGATTTCTATGATGGAGCTGTT 1016
QY 604 GTGGCGGACTCAGTATGCGGCTCTGGCCAGCTGGCAGATGTTGTGGTGGGCTGGAT 663
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1017 GATACAGCATCAGAGAGG---CTAGGCCAATTGACCGATGTGTGCTGCGCTGGAC 1073
QY 664 GACTTTAGGAAGAGTCAGAGAGTGGGGTCTGGCCAGCTATGACTATGTGGGATGGAGC 723
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1074 GATTTACCCAGCACCATGAATACACCGTGTGGCCCGCTATGACTATGTGGGCTGGCGG 1133
QY 724 ACCACAGCTTCTCAGTGGCTATGTGAGATGGAGTTGAGTTGACCGGCTGAGGGCC 783
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1134 AACGAGAGTGCCCAATGGCTACATTTGAGATCATGTTTGAATTTGACCGCATCAGGAAT 1193
QY 784 TTCCAGGCTATCAGGTCCACTGTAAACATGCACACGCTGGGAGCCGCTCTGCTGGC 843
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1194 TTCACTACCATGAGTCCACTGCACACATGTTTGTAAAGTGTGAAGTCTTTAAG 1253
QY 844 GGGGTGAATGTGCTTCGCGGCTGGCCCTGCCATGGCTGGGAGGGAGGCCCATGGCC 903
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1254 GAGGTACAGTGTACTTCCGCTCTGAA---GCCAGTGAAGTGGGAACCTTAATGCCATTTCC 1310
QY 904 CACAACTAGGGGCAACCTGGGGAGCCCGAGCCCGGCTGTCTAGTCCCGCTTGGC 963
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1311 TTCCCGCTTGTCTGGATGACGTCAACCCAGTGTGCTGCGTTCACGGTGGCTCTCCAC 1370
QY 964 GCGCGTGTGGCTCGCTTCTGAGTGGCCGCTTCTCTTTTGGGGGCGCTGTTACTCTTC 1023
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1371 CACCGAATGGCCAGTCCCATCAAGTGTCAATACCATTTTGCAGATACCTGGATGATGTC 1430
QY 1024 AGCGAATCTCTTCACTCTGATGTGGTGAACAATTCCTCTCCGGCACTGGGAGGCACC 1083
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1431 AGTGAGATCACTTCCCAATCAGATGTGCAATGTACAACTCTGAAGCCCTGCCCCACC 1490
QY 1084 TTCCCGCCAGCCCC 1098
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1491 TCTCCTATGGCACCC 1505

```

```

RESULT 9
V55895
ID V55895 standard; DNA; 3120 BP.
AC V55895;
DT 19-NOV-1998 (first entry)
DE Receptor protein tyrosine kinase (PTK) subtype tyro-10 encoding DNA.
KW PTK; receptor; protein tyrosine kinase; brain tissue; ss.
OS Rattus sp.
FH Key Location/Qualifiers
FT CDS 485..3049

```

```

/*tag= a
/product= *PTK subtype tyro-10*

US5811516-A.
22-SEP-1998.
PD 02-JUN-1995; 456647.
PF 15-MAY-1992; US-884486.
PR 02-MAY-1994; US-237401.
PR 02-JUN-1995; US-456647.
PI (SALK ) SALK INST BIOLOGICAL STUDIES.
PI Lai CHC, Lemke GE.
DR WPI: 98-530939/45.
DR P-PSDB: W79152.
PT Receptor protein tyrosine kinase polypeptide, tyro-3 - preferably
    expressed in brain tissue
PS Example 2; Columns 53-60; 46pp; English.
CC This DNA encodes a novel receptor protein tyrosine kinase (PTK)
    polypeptide subtype tyro-10. The invention provides polynucleotide
    sequences encoding novel PTK polypeptide subtypes tyro-1 to tyro-13,
    CC the PTK subtypes are found expressed predominantly in the brain tissue.
SQ Sequence 3120 BP; 760 A; 786 C; 830 G; 744 T;

```

Query Match 26.8%; Score 321; DB 1; Length 3120;

Best Local Similarity 58.4%; Pred. No. 2.2e-67;
Matches 640; Conservative 0; Mismatches 440; Indels 15; Gaps 4;

```

QY 13 AAGGGACATTTTGTATCCTGCCAAGTCCCGCTATGCCCTGGGCATGCAGACCGGACCATC 72
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 548 AAAGCTCAGGTTAATCCAGCCATATGCCGCTATCCTCTGGCATGTCAGAGGCCACATT 607
QY 73 CCAGACAGTGCATATCTCTGTTCCAGCTCCTGTTCAGATTCCACTGCCCGCCGACACAGC 132
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 608 CCAGATGAGGACATCATCAGCCCTCAAGTCAGTGTGTGAGATCCACGCTGCCAAATATGGG 667
QY 133 AGCTTGAGAGAGCAGTGCAGCGGATGGGCTGTGTGCCCGCAGGTCGCTGTTTCCCAAG 192
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 668 AGCTTGAGCTCTCAAGAGAGATGAGGCTGTGTCTCTGAGATTCAGATGCAACCCGAT 727
QY 193 GA---GGAGGAGTACTTGCAGTGTGATCTACAACGACTGCACCTGGTGGCTCTGTGGGC 249
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 728 GACCTGAAGGAATTTCTGCAGATTGACTTGCAGAACCTACACTTTATCACTCTTGTGGGG 787
QY 250 ACCCAGGAGCGCATCGCGGGGCTTGGGAAGGAGTTCTCCCGGAGCTACCGGCTCGCT 309
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 788 ACCCAGGGCGCCATCAGGGGTCATGGCATTGAATTTGCACCCATGTACAAGATCAAC 847
QY 310 TACTCCCGGATGCTCGCGCTGGATGGCTGGAAGACCGCTGGGGTCAGGAGGTGATC 369
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 848 TACAGTCGGATGGCAGTCTCGATCTCTGCGCTAACCGGATGGGAGCAGGTGCTT 907
QY 370 TCAGGCAATGAGGACCTCAGGAGTGGTGTCTGAAGACCTTTGGGCCCCCATGTTGCC 429
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 908 GATGAAACAGTAACCTTATGATGATTTCTTGAAGGACTTGGAGCCACCCATCGTCGCC 967
QY 430 CGACTGTTGCTTCTACCCCCGGCTGACCGGCTCATGAGCTCTGTCTGGGGTAGAG 489
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 968 AGATTGTTTCCGCTTATCCCACTCAGTCCACTCCATGAACGTGTGCATGAGGTTGAG 1027
QY 490 CTCTATGCTGCTCTGGAGGATGGACTCTCTGTCTTACACCGCCCTGTGGGGCAGACA 549
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1028 CTTATGTTGTCTGGCTAGATGGCTGGTATCTCTACATGCTCCAGCTGGACAGCAG 1087
QY 550 ATGTATTTATCTGAGGC-----CGTGTACCTCAACGACTCCACCTATGACGACATACC 603
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1088 TTTGTACTCCCTGGAGCTCCATCATTTATCTGAATGATTTCTGTATGATGAGCTGTT 1147
QY 604 GTGGCGGACTGCAATGATGGGCTCTGGGCCAGCTGGCAGATGGTGTGGGGCTGGAT 663
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1148 GGTGTACAGCATGACTGAAGGG---CTAGGCCAGTTGACTGATGAGTATCCCGGCTGGAT 1204
QY 664 GACTTTAGGAAGAGTCAAGGAGTGGGCTGTGGCCAGGCTATGACTATGTGGGATGGAGC 723
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1205 GATTTACCCAGACCCCATGAATACCAGCTGTGGCTGTGCTATGACTACGTGGGATGGCGG 1264

```

QY 724 AACACAGCTTCCAGTGGCTATGTGGAGATGGAGTTGAGTTGACCGCTGAGGGCC 783
 DB 1265 AACGAAAGTCTACCAACGGTTTCATTGAGATCATGTTTGAATTGACCGAATCAGGAAT 1324
 QY 784 TTCCAGGCTATGAGTCCAGTCTAACACATGCACACCTGGAGCCGCTGCTGCTGGC 843
 DB 1325 TTCTACTACCATGAAGTGCACATGCACAAACATGTTTGTCTAAAGGTGTGAAGATTTTAA 1384
 QY 844 GGGGTGAATGTCTGCTTCCTGGGCTGGCCCTGCGATGGCTGGAGGGAGGCCATGCGC 903
 DB 1385 GAGTCCAGTCTACTTTCGCTCGGA--AGCCAGGAGTGGGAACCCACTGCTCTCTAC 1441
 QY 904 CACAACCTTAGGGGCAACCTGGGGACCCAGAGCCCGGGCTCTCTAGTGGCCCTTGGC 963
 DB 1442 TTTCCTCCTGGTCTGAGCATGTGAACCCAGTGGCCGGTGTTCACGGTGGCCCTCCAC 1501
 QY 964 GCGCGTGGCTGCTTCTGAGTGGCCGCTTCTCTTTCGGGGGCTCTGTTTACTCTTC 1023
 DB 1502 CACCGAATGGCAGTGCCATCAAGTGCCAAATACCAATTTTGGCGACACGTGGATGATGTC 1561
 QY 1024 AGCGAATCTCTCTATCTCTGATGTGGTGAACAATTCCTCTCCGCGACTGGGAGGCACC 1083
 DB 1562 AGCGATACATCTTCCATCAGATGCTGCAATGTATACAACTCTGGAGCCCTTCCCACC 1621
 QY 1084 TTCCCGCCAGCCCC 1098
 DB 1622 TCTCCTATGGCACCC 1636

RESULT 10

V65317 standard; DNA; 3120 BP.

AC V65317;
 DT 22-JAN-1999 (first entry)
 DE Receptor protein tyrosine kinase (PTK) subtype tyro-10 encoding DNA.
 KW PTK; receptor; protein tyrosine kinase; recombinant; grafting;
 OS diagnosis; tumour; skin transplant; connective tissue; tyro-10; ss.
 QS Rattus sp.

FH Key Location/Qualifiers

FT 485..3049

FT CDS /*tag= a

FT /product= "PTK subtype tyro-10"

PN US5837448-A.

PD 17-NOV-1998.

PF 02-MAY-1994; 237401.

PR 15-MAY-1992; US-884486.

PR 02-MAY-1994; US-237401.

PA (SALK.) SALK INST BIOLOGICAL STUDIES.

PI Lai CHC, Lemke GE;

DR WPI; 99-023436/02.

DR P-PSDB; W81409.

PT Nucleic acids encoding protein tyrosine kinase subtypes - for

PT identification of new subtypes and treatment of diseases associated

PT with the kinase

PS Claim 1; Columns 53-58; 47pp; English.

CC This DNA encodes a receptor protein tyrosine kinase (PTK) subtype
 CC tyro-10. The invention provides sequences V65308 to V65313, V65315, and
 CC V65317 to V65319 that encode proteins having a tyrosine kinase domain and
 CC a tissue expression pattern of a receptor PTK subtype selected from
 CC tyro-1, tyro-2, tyro-3, tyro-4, tyro-5, tyro-6, tyro-8, tyro-10, tyro-11,
 CC and tyro-12, respectively. The polynucleotides are useful for the
 CC detection of tyrosine kinase domain sequences and detection of tissue
 CC expression patterns of PTK subtypes. The cDNAs can also be injected into
 CC oocytes, the protein expressed, and expression products screened for
 CC using antibodies against tyrosine kinase epitopes. These subtypes
 CC sequences can be used for the design of oligonucleotides, for use in
 CC amplification reactions to isolate other subtype sequences. These
 CC detection protocols are used in the diagnosis of diseases associated with
 CC (receptor) PTKs. Recombinant vectors expressing the subtypes can be used
 CC to treat related diseases e.g. tumours, by introduction of the vectors
 CC into skin transplants, then grafting these into the connective tissue of
 CC the dermis, thus specifically targeting tumours as the proteins are

CC released from the matrix.

SQ Sequence 3120 BP; 760 A; 786 C; 830 G; 744 T;

Query Match 26.8%; Score 321; DB 1; Length 3120;

Best Local Similarity 58.4%; Pred. No. 2.2e-67;

Matches 640; Conservative 0; Mismatches 440; Indels 15; Gaps 4;

QY 13 AAGGACATTTTATCTCTCCCAAGTGGCGTATGCCCTGGGATGACGAGCCGACATC 72
 DB 548 AAGCTCAGGTTAATCCAGCCATATGCCGTATCTCTGGCATGTGAGGAGCCACAT 607
 QY 73 CCAGACAGTGCATCTCTGCTTCCAGCTCCTGCTCAGATTCACATGCCGCCCCACAC 132
 DB 608 CCAGATGAGACATCAGAGCTCAAGTCAGTGGTCAAGATCCACGGCTGCAAAATATGG 667
 QY 133 AGGTTGGAGAGCAGTACAGGGGATGGGCTGTGTCGCCCGCAGAGGTCGGTGTTCCTCAAG 192
 DB 668 AGGCTGAGCTCTGAAGAAGAGATGGAGCCTGTGTCTGAGATTCCAGTGCAACCCGAT 727
 QY 193 GA---GGAGGAGTACTTGCAGTGGATCTACACAGACTGCACCTGGTCTGTGGGCGC 249
 DB 728 GACCTGAAGGAATTTCTGCAGATTGACTTCCGAACCTACACTTTATCTCTGTGGGG 787
 QY 250 ACCCAGGAGCGATCCGGGGCCCTGGGCAAGAGTTCTCCCGGAGCTACCGGCTCGCT 309
 DB 788 ACCCAGGGGCCATCGAGGGGTCAATGGATTGAATTTGCACCCATGACAAAGATCAAC 847
 QY 310 TACTCCGGGATGTCGCCCTGGATGGGCTGGAAGACCGCTGGGGTCAGGAGGTGATC 369
 DB 848 TACAGTCGGGATGGCAGTCGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 907
 QY 370 TCAGGCAATGAGACCTTGGAGGAGTGGTCTGAAGACCTTGGGGCCCCCATGTTGCC 429
 DB 908 GATGAAACAGTAACCCCTTATGATGATTTCTTGAAGACTTGGAGCCACCCATCGTCGC 967
 QY 430 CGACTGTTTCGCTTCTACCCCGGCTGACCGGTGTCATGAGGCTCTGTCTGGGGTAGAG 489
 DB 968 AGATTGTTGCGCTTATCCAGTCACTGACCTCCATGAACGTGTCATGAGGGTTGAG 1027
 QY 490 CTCTATGGCTGCTCTGGAGGAGTGGACTCTCTTCTTACACCGCCCTGTGGGGCAGACA 549
 DB 1028 CTTTATGTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1087
 QY 550 ATGATTTATCTCAGC-----CGTGTACCTCAAGACTCCACCTATGACGAGATACC 603
 DB 1088 TTTTACTCCTCCGGAGGCTCCATCATTTATCTGAATGATTTCTCTATGATGAGGTGT 1147
 QY 604 GTGGCGGACTGCAGTATGGGGTCTGGGCCAGCTGGCAGATGGTGTGGTGGGGTGGAT 663
 DB 1148 GGTACAGCATGACTGAAGGG---CTAGGCCAGTGTGATGGAGATCTCCGGCTGGAT 1204
 QY 664 GACTTTAGGAAGTACAGGAGTGGCGGTCTGGCCAGGCTATGACTATGTGGGATGGAGC 723
 DB 1205 GATTTTACCAGAGCCCATGAATACCACGTGTGGCCTGGCTATGACTACGTGGGATGG 1264
 QY 724 AACCCAGCTTCTCCAGTGGCTATGTGGAGATGGAGTTGAGTTTACCGGCTGAGGGCC 783
 DB 1265 AACGAAGTGTACCAACCGGTTTCATTTGAGATCATGTTTGAATTTGACCGAATACGAAT 1324
 QY 784 TTCAGGCTATGAGGTCCACTGTAAACATATGCACACCTGGAGCCCGCTGCTGCTGCGC 843
 DB 1325 TTTACTACCATGAGTCCACTGCAACACATGTTTGTCTAAAGGTGTGAAGATTTTAA 1384
 QY 844 GGGGTGAATGTCTGCTTCCGGCTGGCCCTGCCATGGCTGGGAGGGGAGGCCATGGC 903
 DB 1385 GAGGTCCAGTGTACTTTCGCTCGGA--AGCCAGGAGTGGGAACCCACTGCTCTCTAC 1441
 QY 904 CACAACCTTAGGGGCAACCTGGGGACCCAGCCCGGGCTGTCTGAGTGGCCCTTGGC 963
 DB 1442 TTTCCCTGCTCTGAGCATGTGAACCCAGTGGCCCGGTTTGTACGCGTGGCCCTCCAC 1501
 QY 964 GGCCGCTGGCTGCTTCTGCAAGTGGCCCTTCTCTTTTGGGGGGCCCTGTTACTCTTC 1023

Db 1502 CACCGAATGGCCAGTGGCCATCAAGTGCCTAATACATTTGCCGACAGCTGGATGATTC 1561
 QY 1024 AGCGAATCTCTTCATCTCTGATGTGGTGAACAATTCCTCTCGGCACTGGGAGCACC 1083
 Db 1562 AGCGAGATCACTTTCCATCAGATGCTGCAATGTATAACAACACTCTGGAGCCCTTCCACC 1621
 QY 1084 TTCGCCGCCAGCCCG 1098
 Db 1622 TCTCCTATGGCACC 1636

RESULT 11

Q64158
 ID Q64158 standard; DNA; 2128 BP.
 AC Q64158.
 DE 03-FEB-1995 (first entry)
 DE Partial coding sequence of tyrosine kinase receptor protein.
 KW Tyrosine kinase; receptor; proto-oncogene; trk; detection;
 KW diagnosis; antibody; treatment; tumour; antisense; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT cds 1..1953
 FT /*tag= a
 FT /product= Partial sequence of protein tyrosine
 FT kinase.
 PN DE4239817-A.
 PN 01-JUN-1994.
 PD 01-NOV-1992; 239817.
 PF 26-NOV-1992; DE-239817.
 PR (CHEM-) CHEMOTHERAPEUTISCHES FORSCHUNG.
 PA Holtrich U, Ruebsamen-waigmann H, Strebhardt K;
 PI WPI: 94-184380/23.
 DR P-PSDB: R54089.
 DR New protein tyrosine kinase and related nucleic acid - vectors,
 PT transformed cells, etc., useful for diagnosis and treatment of
 PT tumours
 PS Claim 4; Page 8; 9pp; German.
 CC The gene is related to the trk proto-oncogene. Antibodies against
 CC the encoded polypeptide are useful for diagnosis and for the
 CC treatment of tumours. The antibodies may also be radiolabelled or
 CC coupled to a cytotoxin for destruction of cancer cells. Antisense
 CC nucleic acid can be used to inhibit gene expression.
 CC Sequence 2128 BP; 532 A; 556 C; 504 G; 536 T;
 SQ

Query Match 12.0%; Score 143.2; DB 1; Length 2128;
 Best Local Similarity 57.0%; Pred. No. 2.9e-25;
 Matches 303; Conservative 0; Mismatches 223; Indels 6; Gaps 2;

QY 567 CGTGACTCAACGACTCCACCTATGACGACATACCGTGGCGGACTGCGAGTATGGGG 626
 Db 12 CATTTATCTGAATGATTCTGCTATGATGGAGCTTTGGATACAGCATGACAGAAGG-- 70
 QY 627 TCTGGCCAGCTGGCAGATGGTGGTGGGCTGGATGACTTTAGGAGAGTCAAGAGCT 686
 Db 70 -CTAGGCCAATTGACCGATGGTGTCTGGCTGGACGATTTACCCAGACCCCATGAATA 128
 QY 687 GCGGCTGCGCAGGCTATGACTATGTGGGATGGAGCAACACAGCTTCTCCAGTGCTA 746
 Db 129 CCACGTGTGCCCCGCTATGACTATGTGGCTGGCGGACGAGAGTGGCCACCATGCTA 188
 QY 747 TGTGAGATGAGGTTGAGTTGACCGCTGAGGCGCTTCCAGGCTATGCAAGTCCACTG 806
 Db 189 CATTGAGATCATGTTGAATTTGACCGCATCAGGAATTTCACTACCATGAAGTCCACTG 248
 QY 807 TAACACATGCACGCTGGGAGCCGCTGCTCCGCGGCTGGATGCTCCGCGG 866
 Db 249 CAACAACATGTTTGAAGGTGTGAAGATCTTTAAGGAGGTACAGTCTACTTCCGCTC 308
 QY 867 TGGCCCTGCGATGGCTGGGAGGGGAGCCCATGCGCCACACACCTAGGGGGCAACCTGGG 926
 Db 309 TGAA--GCCAGTGAGTGGGAACCTAATGCCATTTCCCTTCCCTTGTCTGGATGAGCT 365

QY 927 GGACCCAGAGCCCGCGGTGTCTCAGTCCCTTGGCGGCGGTGTGGCTCCTTTCTGCA 986
 Db 366 CAACCCAGTGTCTGGTTGTACGGTCCCTCTCCACCACCAATGGCCAGTGCATCAA 425
 QY 987 GTGCGGCTTCTCTTTGGGGGCGCTGTACTTTCAGCGAAATCTCTTCATCTCTGA 1046
 Db 426 GTGTCAATACCATTTTTCAGATACCTGGATGATGTTCAGTGGATCACCCTTCCAATCAGA 485
 QY 1047 TGTGTGAACATTCCTCTCGGCACTGGGAGGACCTTCCCGCCAGCCGCC 1098
 Db 486 TCTGCAATGTACAACTCTGAAGCCCTGCCACCTCTCTCTATGGCACCC 537

RESULT 12

V23288
 ID V23288 standard; DNA; 4451 BP.
 AC V23288;
 DT 17-AUG-1998 (first entry)
 DE Synthetic human Factor-VIII gene lacking central B domain.
 KW Factor-VIII; blood clotting; human; synthetic gene; codon usage; ss.
 OS Homo sapiens.
 OS Synthetic.
 PN WO9812207-A1.
 PD 26-MAR-1998.
 PF 18-SEP-1997; U16639.
 PR 20-SEP-1996; US-717294.
 PA (GEO) GEN HOSPITAL CORP.
 PI Haas J, Seed B;
 DR WPI: 98-217200/19.
 DR New synthetic eukaryotic gene(s) - in which non-preferred or less
 PT preferred codon(s) are replaced to provide high level expression in
 PT mammalian cell(s)
 PT Claim 20; Fig 13; 92pp; English.
 CC This synthetic gene codes for a human Factor-VIII protein that
 CC lacks the central B domain (amino acids 760-1639) of the native
 CC protein. In the synthetic gene, non-preferred or less preferred
 CC codons of the native gene (see V23339) are replaced by codons
 CC favored by highly expressed human genes to provide high-level
 CC expression in mammalian cells. The synthetic gene was assembled
 CC from 29 pairs of oligonucleotides (see V23340-97) which served as
 CC PCR templates. Synthetic genes of the invention (see also
 CC V23289-91) are used for production of recombinant proteins in
 CC mammalian cells at levels of at least 500% of those obtained using
 CC the natural genes. They can also be used in gene therapy. An
 CC expression vector comprising a synthetic gene and a mammalian cell
 CC harboring a synthetic gene are also claimed.
 CC Sequence 4451 BP; 972 A; 1568 C; 1234 G; 677 T;
 SQ

Query Match 6.2%; Score 73.8; DB 1; Length 4451;
 Best Local Similarity 50.1%; Pred. No. 9.8e-09;
 Matches 211; Conservative 0; Mismatches 207; Indels 3; Gaps 1;

QY 33 CAAGTCCCGCTATGCCCTGGGCATGCAGACCGGACCATCCAGACAGTACATCTCTGC 92
 Db 3959 CAGCTCAGCATGCCCTGGGCATGGAGAGCAAGGCCATCAGCGACGCCCATCACCGC 4018
 QY 93 TTCACGCTCTGTGTCAAGTTCCACTGCCCGCCGCCACAGCAGGTTGGAGAGCATGACGG 152
 Db 4019 CTCACGCTACTTCAACCAACATGTCGCCACCTGGAGCCGCCAGAGCCGCTGCACCT 4078
 QY 153 GGATGGGCGCTGTGCGCCGCGAGGTGCGTGTTCCTCCAGGAGGAGGATCTTCAGGT 212
 Db 4079 GCAGGGCCGCGACAGCAGCGCTGGCGCCCGCCAGGTGAACACCCCAAGAGTGTGCTCAGST 4138
 QY 213 GGATCTACACGACTGCACCTGTGTGCTGTGGGACCCAGGAGCGCATGCCGGGG 272
 Db 4139 GGAATCCAGAAACCATGAAGTGAAGTGGGTGAGTACCCAGGCGGTCAAGAGCTG-- 4197
 QY 273 CTTGGGCAAGGAGTTCTCCCGGAGCTACCGGTGCGGTACTCCCGGGAGTGTGCGCGCTG 332
 Db 4197 -CTGACGAGCATGTACGTGAAGGAGTTCCTGATCAGCAGCAGCAGGACGCGCCACCATGTG 4255

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 4, 1999, 07:31:19 ; Search time 1266.11 Seconds
(without alignments)
3006.714 Million cell updates/sec

Title: US-08-170-558-7
Perfect score: 1197
Sequence: 1 GATCGTACATGAGGACAC.....CCGAGGGAGCCGACGCGC 1197

Scoring table: IDENTITY_NUC

Searched: 679419 seqs, 1590154680 residues

Database : GenEmbl.*
1: gb_ba1.*
2: gb_ba2.*
3: gb_om.*
4: gb_ov.*
5: gb_pat.*
6: gb_ph.*
7: gb_pl1.*
8: gb_pl2.*
9: gb_pr1.*
10: gb_pr2.*
11: gb_pr3.*
12: gb_ro.*
13: gb_st.*
14: gb_sts.*
15: gb_sy.*
16: gb_un.*
17: gb_vi.*
18: em_fun.*
19: em_htg.*
20: em_hum1.*
21: em_hum2.*
22: em_in.*
23: em_om.*
24: em_or.*
25: em_ov.*
26: em_pat.*
27: em_ph.*
28: em_pl.*
29: em_ro.*
30: em_sts.*
31: em_sy.*
32: em_un.*
33: em_vi.*
34: gb_htg1.*
35: gb_htg2.*
36: gb_in1.*
37: gb_in2.*
38: em_ba1.*
39: em_ba2.*
40: em_hum3.*
41: em_hum4.*
42: gb_pr4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match Length	DB ID	Description
1	1197	100.0	3637 5 180845	180845 Sequence 3

ALIGNMENTS

RESULT	1	LOCUS	180845	Sequence 3	3637 bp	DNA	PAT	20-MAR-1998
DEFINITION				Sequence 3 from patent US 5709858.				
ACCESSION			180845					
NID			93209135					
VERSION			180845.1	GI:3209135				
KEYWORDS								
SOURCE			Unknown.					
ORGANISM			Unclassified.					
REFERENCE			1 (bases 1 to 3637)					
AUTHORS			Godowski, P., J., Mark, M.R. and Scadden, D.T.					
TITLE			Antibodies specific for Rse receptor protein tyrosine kinase					
JOURNAL			Patent: US 5709858-A 3 20-JAN-1998;					
FEATURES			Location/Qualifiers					
			1..3637					
			/organism="unknown"					
			BASE COUNT 721 a 1094 c 1079 g 743 t					
			ORIGIN					

Query Match 100.0%; Score 1197; DB 5; Length 3637;
Best Local Similarity 100.0%; Pred. No. 6.7e-249;

Matches 1197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATCCTGACATGAAGGACATTTTATCTTCCCAAGTGGCGCTATGCTTGGCGATGCGAG 60
 Db 136 GATCCTGACATGAAGGACATTTTATCTTCCCAAGTGGCGCTATGCTTGGCGATGCGAG 195
 QY 61 GACCGGACCATCCAGACAGTACATCTCTGCTTCCAGTCTCTGCTGAGATTCACATGCC 120
 Db 196 GACCGGACCATCCAGACAGTACATCTCTGCTTCCAGTCTCTGCTGAGATTCACATGCC 255
 QY 121 GCCCGCACAGCAGTGGAGAGCAGTACGGGATGGGGCTGGTGGCGGAGGCTG 180
 Db 256 GCCCGCACAGCAGTGGAGAGCAGTACGGGATGGGGCTGGTGGCGGAGGCTG 315
 QY 181 GTGTTTCCAGGAGGAGGAGTACTTCCAGTGGATCTACACGACTGACCTGGTGGCT 240
 Db 316 GTGTTTCCAGGAGGAGGAGTACTTCCAGTGGATCTACACGACTGACCTGGTGGCT 375
 QY 241 CTGTTGGGACCCAGGACGSCATCCCGGGGGCTGGGCAAGGAGTCTCCCGGAGCTAC 300
 Db 376 CTGTTGGGACCCAGGACGSCATCCCGGGGGCTGGGCAAGGAGTCTCCCGGAGCTAC 435
 QY 301 CGGTGCTTACTCCCGGATGCTGGCGCTGGATGGGCTGGAGGACCGCTGGGCTAG 360
 Db 436 CGGTGCTTACTCCCGGATGCTGGCGCTGGATGGGCTGGAGGACCGCTGGGCTAG 495
 QY 361 GAGGTGATCTCAGCAATGAGGACCTGAGGAGTGGTCTGAAGGACCTTGGGCGCCGCC 420
 Db 496 GAGGTGATCTCAGCAATGAGGACCTGAGGAGTGGTCTGAAGGACCTTGGGCGCCGCC 555
 QY 421 ATGTTGCCCGACTGGTTCCTTACCCCGGGCTGACCGGGCTCATGAGCGTCTGCTG 480
 Db 556 ATGTTGCCCGACTGGTTCCTTACCCCGGGCTGACCGGGCTCATGAGCGTCTGCTG 615
 QY 481 CGGTGAGCTCTATGCTGCTCTGAGGAGTGGATCTCTGCTTACACCGCCCTG 540
 Db 616 CGGTGAGCTCTATGCTGCTCTGAGGAGTGGATCTCTGCTTACACCGCCCTG 675
 QY 541 GGGCAGACAATGATTTATCTGAGGCGCTGTACCTCAACGACTCCACTATGACGGACAT 600
 Db 676 GGGCAGACAATGATTTATCTGAGGCGCTGTACCTCAACGACTCCACTATGACGGACAT 735
 QY 601 ACCGTGGCGGAGTGCAGTATGGGGCTGGGCGAGTGGGAGTGGTGGGGCTG 660
 Db 736 ACCGTGGCGGAGTGCAGTATGGGGCTGGGCGAGTGGGAGTGGTGGGGCTG 795
 QY 661 GATGACTTTAGAGAGTGCAGAGTGGGCTGCGGCTGCGCAGCTATGACTATGTTGGGATG 720
 Db 796 GATGACTTTAGAGAGTGCAGAGTGGGCTGCGGCTGCGCAGCTATGACTATGTTGGGATG 855
 QY 721 AGCAACACAGCTTCTCCAGTGGCTATGTTGAGATGGAGTTTGGATTTGACCGGCTGAGG 780
 Db 856 AGCAACACAGCTTCTCCAGTGGCTATGTTGAGATGGAGTTTGGATTTGACCGGCTGAGG 915
 QY 781 GCTTTCCAGGATGACAGTCCACTGTAAACATGACACGCTGGAGGCGCTCTGCT 840
 Db 916 GCTTTCCAGGATGACAGTCCACTGTAAACATGACACGCTGGAGGCGCTCTGCT 975
 QY 841 GCGGGGTGGAATGCTGCTTCCGGCTGCGCTGCGCATGGCTGGGAGGGGAGCCCATG 900
 Db 976 GCGGGGTGGAATGCTGCTTCCGGCTGCGCTGCGCATGGCTGGGAGGGGAGCCCATG 1035
 QY 901 GCGCAACACCTAGGGGGAACCTGGGGGACCCAGAGCCCGGGCTGTCTAGTGGCCCTT 960
 Db 1036 GCGCAACACCTAGGGGGAACCTGGGGGACCCAGAGCCCGGGCTGTCTAGTGGCCCTT 1095
 QY 961 GCGGCGGTGGTGGCTGCTTCTGAGTGGCGCTTCTTCTTGGGCGCCCTGTTACTC 1020
 Db 1096 GCGGCGGTGGTGGCTGCTTCTGAGTGGCGCTTCTTCTTGGGCGCCCTGTTACTC 1155
 QY 1021 TTCAGGAAATCTCTCTATCTGATGTTGTAACAAATCTCTCCGGCACTGGGAGGC 1080
 Db 1156 TTCAGGAAATCTCTCTATCTGATGTTGTAACAAATCTCTCCGGCACTGGGAGGC 1215

QY 1081 ACCTTCCCGCAGCCCTGCTGTCGCCCTGGCCACACTTCCACCACTTTCAGCAGCTTG 1140
 Db 1216 ACCTTCCCGCAGCCCTGCTGTCGCCCTGGCCACACTTCCACCACTTTCAGCAGCTTG 1275
 QY 1141 GAGCTGGAGCCCGCAGGCGCAGCAGCCCGTGGCCAAAGCCCGAGGGAGCCGACGCC 1197
 Db 1276 GAGCTGGAGCCCGCAGGCGCAGCAGCCCGTGGCCAAAGCCCGAGGGAGCCGACGCC 1332

RESULT 2

LOCUS I80847 1197 bp DNA PAT 20-MAR-1998
 DEFINITION Sequence 7 from patent US 5709858.
 ACCESSION I80847
 NID 93209137
 VERSION I80847.1 GI:3209137
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassefied.
 REFERENCE 1 (bases 1 to 1197)
 AUTHORS Godowski, P.J., Mark, M.R. and Scadden, D.T.
 TITLE Antibodies specific for Rse receptor protein tyrosine kinase
 JOURNAL Patent: US 5709858-A 7 20-JAN-1998;
 FEATURES
 1. Location/Qualifiers
 1. 1197
 source /organism="unknown"
 BASE COUNT 199 a 351 c 407 g 240 t
 ORIGIN

Query Match 100.0%; Score 1197; DB 5; Length 1197;
 Best Local Similarity 100.0%; Pred. No. 7.6e-249;
 Matches 1197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATGTGACATGAAGGACATTTTATCTTCCCAAGTGGCGCTATGCTTGGCGATGCGAG 60
 Db 1 GATGTGACATGAAGGACATTTTATCTTCCCAAGTGGCGCTATGCTTGGCGATGCGAG 60
 QY 61 GACCGGACCATCCAGACAGTACATCTCTGCTTCCAGTCTCTGCTGAGATTCACATGCC 120
 Db 61 GACCGGACCATCCAGACAGTACATCTCTGCTTCCAGTCTCTGCTGAGATTCACATGCC 120
 QY 121 GCCCGCACAGCAGTGGAGAGCAGTACGGGATGGGGCTGGTGGCGGAGGCTG 180
 Db 121 GCCCGCACAGCAGTGGAGAGCAGTACGGGATGGGGCTGGTGGCGGAGGCTG 180
 QY 181 GTGTTTCCCAAGGAGGAGGAGTACTTGCAGTGGATCTACACGACTGCACCTGGTGGCT 240
 Db 181 GTGTTTCCCAAGGAGGAGGAGTACTTGCAGTGGATCTACACGACTGCACCTGGTGGCT 240
 QY 241 CTGTTGGGACCCAGGACGCGCATCCCGGGGCTGGGCAAGGAGTCTCCCGGAGCTAC 300
 Db 241 CTGTTGGGACCCAGGACGCGCATCCCGGGGCTGGGCAAGGAGTCTCCCGGAGCTAC 300
 QY 301 CGGCTGCTTACTTCCCGGATGGTCCCGCTGGATGGCTGGAGGACCGCTGGGGCTGAG 360
 Db 301 CGGCTGCTTACTTCCCGGATGGTCCCGCTGGATGGCTGGAGGACCGCTGGGGCTGAG 360
 QY 361 GAGGTGATCTCAGGCAATGAGGACCGCTGAGGAGTGGTGGTGAAGGACCTTGGGCGCCGCC 420
 Db 361 GAGGTGATCTCAGGCAATGAGGACCGCTGAGGAGTGGTGGTGAAGGACCTTGGGCGCCGCC 420
 QY 421 ATGTTTCCCGACTGGTTCGCTTCTACCCCGGGCTGACCGGGCTCATGAGCGTCTGCTG 480
 Db 421 ATGTTTCCCGACTGGTTCGCTTCTACCCCGGGCTGACCGGGCTCATGAGCGTCTGCTG 480
 QY 481 CGGCTGAGCTCTATGCTGCTCTGGAGGATGAGTCTCTCTTTACACCGCCCTG 540
 Db 481 CGGCTGAGCTCTATGCTGCTCTGGAGGATGAGTCTCTCTTTACACCGCCCTG 540
 QY 541 GGCACACAAATGATTTATCTGTAGGCGCGTGTACCTCAAGGACTCCACCTATGACGACAT 600

```

Db 541 GGGAGACATGATTTATCTGAGCCGCTGATACCTCAACGACTCCACCTATGACGACAT 600
Qy 601 ACCGTGGCGGACTGCAGTATGGGGTCTTGGCCAGCTGGCAGATGCTGTGGGGCTG 660
Db 601 ACCGTGGCGGACTGCAGTATGGGGTCTTGGCCAGCTGGCAGATGCTGTGGGGCTG 660
Qy 661 GATGACATTTAGGAGATCAGGAGCTCGGGTCTGGCCAGCTGATGACTATGATGGATGG 720
Db 661 GATGACATTTAGGAGATCAGGAGCTCGGGTCTGGCCAGCTGATGACTATGATGGATGG 720
Qy 721 AGCAACACAGCTTCTCAGTGGCTATGAGATGAGTTCAGTTCGACCGCTGAGG 780
Db 721 AGCAACACAGCTTCTCAGTGGCTATGAGATGAGTTCAGTTCGACCGCTGAGG 780
Qy 781 GCCTTCCAGGCTATGAGGCTCACTTAACAACATGACACGCTGGGAGCCCTGCTGCCCT 840
Db 781 GCCTTCCAGGCTATGAGGCTCACTTAACAACATGACACGCTGGGAGCCCTGCTGCCCT 840
Qy 841 GCGGGGTGGAATCTCGCTTCCGGCTGGCCCTGCCATGGCTGGGAGGGAGCCCATG 900
Db 841 GCGGGGTGGAATCTCGCTTCCGGCTGGCCCTGCCATGGCTGGGAGGGAGCCCATG 900
Qy 901 CGCCACACCTAGGCGGCACTCGGGGACCCAGAGCCCGGCTGCTCAGTGGCCCTT 960
Db 901 CGCCACACCTAGGCGGCACTCGGGGACCCAGAGCCCGGCTGCTCAGTGGCCCTT 960
Qy 961 GCGGCGCTGTGGCTTCTTCACTGCGCTTCTTCTTGGGGGCGCTTGGTACTC 1020
Db 961 GCGGCGCTGTGGCTTCTTCACTGCGCTTCTTCTTGGGGGCGCTTGGTACTC 1020
Qy 1021 TTCAGGAAATCTCCTTCACTCTGATGTGGTGAACAATCTCTTCCGGCACTGGGAGGC 1080
Db 1021 TTCAGGAAATCTCCTTCACTCTGATGTGGTGAACAATCTCTTCCGGCACTGGGAGGC 1080
Qy 1081 ACCTTCCGCGAGCCCTTGGTGGCCGCTGGCCAGCCCGGAGGAGCCGAGCGCTG 1140
Db 1081 ACCTTCCGCGAGCCCTTGGTGGCCGCTGGCCAGCCCGGAGGAGCCGAGCGCTG 1140
Qy 1141 GAGCTGGAGCCAGAGGCGGAGCCGCTGGCCAGCCCGGAGGAGCCGAGCGCTG 1197
Db 1141 GAGCTGGAGCCAGAGGCGGAGCCGCTGGCCAGCCCGGAGGAGCCGAGCGCTG 1197

RESULT 3
HUMCAK HUMCAK 3803 bp mRNA PRI 11-MAY-1994
LOCUS Homo sapiens tyrosine protein kinase (CAK) gene, complete cds.
ACCESSION L20817
NID g306474
VERSION L20817.1 GI:306474
KEYWORDS A5-antigen; factor V; factor VIII; tyrosine protein kinase.
SOURCE Homo sapiens fetus liver cDNA to mRNA.
ORGANISM Homo sapiens

REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 3803)
Perez, J.L., Shen, X., Finkernagel, S., Sciorra, L., Jenkins, N.A.,
Gilbert, D.J., Copeland, N.G. and Wong, T.W.
Identification and chromosomal mapping of a receptor tyrosine
kinase with a putative phospholipid binding sequence in its
ectodomain
Oncogene 9, 211-219 (1994)
MEDLINE 94134417
REFERENCE 2 (bases 1 to 3803)
Johnson, J.D.
Direct Submission
TITLE Submitted (17-FEB-1993) J.D. Johnson
JOURNAL 3 (sites)
REFERENCE Johnson, J.D., Edman, J.C. and Rutter, W.J.
TITLE A receptor tyrosine kinase found in breast carcinoma cells has an
extracellular discoidin I-like domain
Proc. Natl. Acad. Sci. U.S.A. 90, 5677-5681 (1993)
JOURNAL

```

```

MEDLINE 93296201
REFERENCE.. 4 (sites)
AUTHORS Wong, T.
TITLE Direct Submission
JOURNAL Submitted (06-JUL-1993) T. Wong
FEATURES
    Location/Qualifiers
        1..3803
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /dev_stage="fetus"
            /tissue_type="liver"
            /map="chromosome 6"
            /gene="CAK"
            1..2934
            1..192
            /gene="CAK"
            193..246
            /gene="CAK"
            193..2934
            /gene="CAK"
            /note="N-linked glycosylation sites: bps 829-831, 976-978,
            1306-1308, 1378-1380; domains: transmembrane-bps.
            1441-1521, kinase-bps. 2041-2877; furin cleavage site:
            bps. 1102-1113"
            /codon_start=1
            /product="tyrosine protein kinase"
            /protein_id="AAAI8019.1"
            /db_xref="PID:g306475"
            /db_xref="GI:306475"
            /translation="MGPEALSSLLLLLLVASGDADMKGHFDPAKCRYALGMDRTIIPD
            SDISASSWSSTAAHRSLRSSDGDGAWCPAGSVFPKEEYLQVDLQRLHLVALVGT
            QRRHAGGLGKFEFSRSLRYSRDRGRWGMKDRWQEVISGNEDEPGVLDKLGPPMV
            ARLVRFYPRADRVMSVCLRVLYGLWRDGLLSYAPVQGTMYLSEAVYLDSTYDGH
            TVGGLGYGLQGLADVVDLDRKSOELRWKVPDYDVVGNHSSFSVGVYEMEPEDFR
            LRAQAMQVHCNMHTLGARLPGGVECFRFRGPMANEGEPHNLGNLGDPRARAV
            SVPLGVRVARELQCFRLFAGFWLLFSEISFISDVVNSSPALGGTFFPAWMPGPPPP
            TNSFLEPRGQQPQAKASGPTAILGLVAIILLILLIALLMLRLHWRLLSKA
            ERRYLEELTVHLSVPGDITLILNNRPGPREPPYQEPGRNPPHSPACPNWNSALLL
            SNPAYRLLLATYARPPRGPPPTPAWAKPTNTQAYSGDYMEPEKPGAPLLPPPPNSV
            PHYAEADITVLTGGTVAVPALPGAVDGPVRYDFPRSLRFRFKELGQGFQGV
            HLCVDSQDLVSLDFPLNVRKHPHLLVAVKILRPDATKNARNDFLKVKIMSLKDP
            NIIRLLGVCVQDDPLCMITDMENGLNQFLSAHQEDKRAEAGPQGGQAAGGTISF
            PMLLHVAQAAGSMWECILMGKFTTASDVWAFVTLWEVLMCLRAQPFQGLTDEQVIER
            VQGRAVLPIRMAWECILMGKFTTASDVWAFVTLWEVLMCLRAQPFQGLTDEQVIER
            AGEFFRDQGRGVLYLSRPPACPGQGLYELMLRCWSRESQRPPFSQLHFLAEDALNTV"
BASE COUNT 721 a 1184 c 1123 g 775 t
ORIGIN

Query Match 99.9%; Score 1195.4; DB 9; Length 3803;
Best Local Similarity 99.9%; Pred. No. 1.5e-248;
Matches 1196; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GATGCTGACATGAAGGACATTTTATCTGCCAAGTCCGCTATGCTCCCTGGGCATGCAG 60
Db 247 GATGCTGACATGAAGGACATTTTATCTGCCAAGTCCGCTATGCTCCCTGGGCATGCAG 306
Qy 61 GACCGACCATCCAGACAGTACATCTCTCCAGCTCTGCTGCTGCTGCTGCTGCTGCTG 120
Db 307 GACCGACCATCCAGACAGTACATCTCTCCAGCTCTGCTGCTGCTGCTGCTGCTGCTG 366
Qy 121 GCGCGCCACAGCAGGTTGGAGACAGTACGCGGAGTGGGGCTGCTGCTGCTGCTGCTGCTG 180
Db 367 GCGCGCCACAGCAGGTTGGAGACAGTACGCGGAGTGGGGCTGCTGCTGCTGCTGCTGCTG 426
Qy 181 GTGTTTCCCAAGAGAGAGGAGTACTTTCAGGTGGATCTACACAGCTGCACCTGCTGCTG 240
Db 427 GTGTTTCCCAAGAGAGAGGAGTACTTTCAGGTGGATCTACACAGCTGCACCTGCTGCTG 486
Qy 241 CTGGTGGGACCCAGGAGGAGGATGCTGGGGGCGCTTGGCAAGAGGATTTCTCCGGAGCTAC 300
Db 487 CTGGTGGGACCCAGGAGGAGGATGCTGGGGGCGCTTGGCAAGAGGATTTCTCCGGAGCTAC 546

```

301 CGGTCGGTCTACTCCGGGATGCTCGCGCTGGATGGCTGAAGGACCGCTGGGTCAG 360
Db 547 CGGTCGGTCTACTCCGGGATGCTCGCGCTGGATGGCTGAAGGACCGCTGGGTCAG 606
QY 361 GAGGTGATCTCAGGAATGAGGACCTGAGGAGTGGTCTGAAGGACCTTGGGCCCCC 420
Db 607 GAGGTGATCTCAGGAATGAGGACCTGAGGAGTGGTCTGAAGGACCTTGGGCCCCC 666
QY 421 ATGTTGCCCGACTGGTTCGCTTACCCCCGGGCTGACCGGGTATGAGGCTGTGCTG 480
Db 667 ATGTTGCCCGACTGGTTCGCTTACCCCCGGGCTGACCGGGTATGAGGCTGTGCTG 726
QY 481 CGGTAGAGCTCTATGGCTGCTCTGGAGGATGACCTCTTACACCGCCCTGTG 540
Db 727 CGGTAGAGCTCTATGGCTGCTCTGGAGGATGACCTCTTACACCGCCCTGTG 786
QY 541 GGGCAGACAATGATTTATCTCAGGCGCTGTACCTCAACGACTCCACCTATGACGGACAT 600
Db 787 GGGCAGACAATGATTTATCTCAGGCGCTGTACCTCAACGACTCCACCTATGACGGACAT 846
QY 601 ACCGTGGCGGACTCAGATATGGGGTCTGGGCCAGCTGGCAGATGTGTGGGGCTG 660
Db 847 ACCGTGGCGGACTCAGATATGGGGTCTGGGCCAGCTGGCAGATGTGTGGGGCTG 906
QY 661 GATGACTTTAGGAAGAGTCAGAGCTCGGGTCTGGCCAGCTATGACTATGGGATGG 720
Db 907 GATGACTTTAGGAAGAGTCAGAGCTCGGGTCTGGCCAGCTATGACTATGGGATGG 966
QY 721 AGCAACACAGCTTCTCCAGTGGCTATGTGGAGATGGATGTAGTTGACCGCTGAGG 780
Db 967 AGCAACACAGCTTCTCCAGTGGCTATGTGGAGATGGATGTAGTTGACCGCTGAGG 1026
QY 781 GCTTCCAGGCTATCAGGTCCACTGTAAACATGTCACAGCTGGGAGCCCTGTGCT 840
Db 1027 GCTTCCAGGCTATCAGGTCCACTGTAAACATGTCACAGCTGGGAGCCCTGTGCT 1086
QY 841 GCGGGGTGGAATGTCGCTTCGGGCTGGCCCTGCATGGCTGGGAGGGAGCCCATG 900
Db 1087 GCGGGGTGGAATGTCGCTTCGGGCTGGCCCTGCATGGCTGGGAGGGAGCCCATG 1146
QY 901 CGCCACAACCTAGGGGGAACCTGGGGGACCCCGAGAGCCCGGCTGTCTCAGTGGCCCTT 960
Db 1147 CGCCACAACCTAGGGGGAACCTGGGGGACCCCGAGAGCCCGGCTGTCTCAGTGGCCCTT 1206
QY 961 GCGGCGGTGGCTGCTGCTTCAGTGGCGCTTCCTCTTTGGGGGCGCTGTGTTACT 1020
Db 1207 GCGGCGGTGGCTGCTGCTTCAGTGGCGCTTCCTCTTTGGGGGCGCTGTGTTACT 1266
QY 1021 TTCAGGAAATCTCCTTCATCTCTGATGTTGACAAATTCCTTCCGGCACTGGGAGGC 1080
Db 1267 TTCAGGAAATCTCCTTCATCTCTGATGTTGACAAATTCCTTCCGGCACTGGGAGGC 1326
QY 1081 ACCTTCCGCGACCCCTGTGGCGCTGGCCACCTCCACCACTCCAGCAGCTTG 1140
Db 1327 ACCTTCCGCGACCCCTGTGGCGCTGGCCACCTCCACCACTCCAGCAGCTTG 1386
QY 1141 GAGCTGGAGCCAGAGGCGCAGAGCCCTGGCCAGCCCGGAGGAGCCCGACCGCC 1197
Db 1387 GAGCTGGAGCCAGAGGCGCAGAGCCCTGGCCAGCCCGGAGGAGCCCGACCGCC 1443

RESULT 4

HSRETYK1 HSRETYK1 3841 bp mRNA PRI 20-APR-1995
LOCUS H.sapiens EDDR1 gene for receptor tyrosine kinase.
DEFINITION
ACCESSION Z29093
NID 9732799
VERSION Z29093.1 GI:732799
KEYWORDS receptor tyrosine kinase.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 3841)
Laval,S., Butler,R., Shelling,A.N., Hanby,A.M., Poulsom,R. and Ganesan,T.S.
TITLE Isolation and characterization of an epithelial-specific receptor tyrosine kinase from an ovarian cancer cell line
JOURNAL Cell Growth Differ. 5 (11), 1173-1183 (1994)
MEDLINE 95151638
REFERENCE 2 (bases 1 to 3841)
Shelling,A.N., Butler,R., Jones,T., Laval,S., Boyle,J.M. and Ganesan,T.S.
TITLE Localization of an epithelial-specific receptor kinase (EDDR1) to chromosome 6q16
JOURNAL Genomics 25 (2), 584-587 (1995)
MEDLINE 95309932
REFERENCE 3 (bases 1 to 3841)
Kedinger,C.
TITLE Direct Submission
JOURNAL Submitted (17-DEC-1993) Claude Kedinger, CNRS Laboratoire de genetique moleculaire-U184, INSERM, 11, rue Humann, Strasbourg, Alsace, 67085 cedex, FRANCE
FEATURES
source
1..3841
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/clone="RTK6"
/dev_stage="Adult"
/tissue_type="Ovary"
/cell_type="Epithelial cell"
/cell_line="SKOV-3"
/clone_lib="pks"
336..423
/note="putative signal peptide"
337..2967
/gene="EDDR1"
337..2967
/gene="EDDR1"
/standard_name="epithelial discoidin domain receptor"
/product="receptor tyrosine kinase"
/protein_id="CAA82335.1"
/db_xref="PID:g732800"
/db_xref="GI:732800"
/db_xref="SWISS-PROT:O08345"
/translation="MGPEALSLLLLLLVASGDADMGHFDPAKRYALGMQDRITPD SDIASSSSDSTAARHSRLSSDGDGAWCPAGSVFPEEYLOVDLQRLHLVALVGT QGRHAGLGKESRYRLYSRDRGRWGWKDRGOEVIQSGNEDPGVLDLQPPMV ARLYFYPRADRVMSVCLRVLYGCLWRDGLLSYAPVQWTYLSAVYLNDSYDGH LRAFOAMQVHCNNMHTLGAFLPGVVECFRRGPAMAEPEMHNGLNGLDGPRAV TNFSLLEPRGQOPVAKAGSPATILIGCLVAIILLILLIALLMLRHLWRLLSKA ERRLVEELTVHLSVPGDTILINRPGPREPPYOEPRGNP9HSPACVPNGSAYSG DYMEPEKCAPLLPPPPQNSVPHYAADIIVTQGTGNTYAVLPPGAVDGPFRV DFRSRLEKELGEGVEHLCVDSQDLVSLDPFLNVKRGDLPPLVAVKILRPA TKNARDFLEKVKIMSLKDNIIIRLLGCVQDDPLCMITDYMNGDLNQFLSAHLE DKAAEAGPDGQAQGTPTSYPMLLHVAQAISGMRYLATLNFVRDLATRNCLVGEN FTIKADFMSRNLVAGDYRYVQGRAVLPIRMAWECILMGKETTASDVWAFVTLWE VMLCRAOPFGOLTDEQVIEAGFEFRDQGRVLSRPPACQGLYELMLRCWRESE ORPPFSQLHRLAEDALNTV"
polyA_signal 3807
polyA_site 3828
BASE COUNT 750 a 1123 c 1126 g 842 t
ORIGIN
Query Match 99.9%; Score 1195.4; DB 10; Length 3841;
Best Local Similarity 99.9%; Pred. No. 1.5e-248;
Matches 1196; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GATGCTGACATGAAGGACATTTTGTATCTCCGCAAGTCCCGCTATGCTGGGATGAG 60
Db 391 GATGCTGACATGAAGGACATTTTGTATCTCCGCAAGTCCCGCTATGCTGGGATGAG 450

QY 61 GACCGGACCATCCAGACAGTACATCTCTGCTTCAGCTCCTGTCAGATTCACATGCC 120
DB 451 GACCGGACCATCCAGACAGTACATCTCTGCTTCAGCTCCTGTCAGATTCACATGCC 510
QY 121 GCGCGCCACAGAGTTGGAGAGCAGTACGCGGATGGGCTGGTCCCGCAGGTCG 180
DB 511 GCGCGCCACAGAGTTGGAGAGCAGTACGCGGATGGGCTGGTCCCGCAGGTCG 570
QY 181 GTGTTTCCCAAGAGAGGAGTACTTGACAGTGGATCTACACGACTGCACCTGTGCT 240
DB 571 GTGTTTCCCAAGAGAGGAGTACTTGACAGTGGATCTACACGACTGCACCTGTGCT 630
QY 241 CTGGTGGGACCCAGGAGCAGTACGCGGAGGCTGGGCAAGAGTCTCCCGAGCTAC 300
DB 631 CTGGTGGGACCCAGGAGCAGTACGCGGAGGCTGGGCAAGAGTCTCCCGAGCTAC 690
QY 301 CGGCTGCGTACTCCCGGATGGTCCGCTGGATGGGCTGGAAGCAGCTGGGGTCAG 360
DB 691 CGGCTGCGTACTCCCGGATGGTCCGCTGGATGGGCTGGAAGCAGCTGGGGTCAG 750
QY 361 GAGGTATCTCAGGCAATGAGGACCTCTGAGGAGTGGTCTGAGGAGCTTTGGGCCCCC 420
DB 751 GAGGTATCTCAGGCAATGAGGACCTCTGAGGAGTGGTCTGAGGAGCTTTGGGCCCCC 810
QY 421 ATGTTTCCCGACTGGTTCGCTTACCCCGGCTGACCGGCTCATGAGCTCTGCTG 480
DB 811 ATGTTTCCCGACTGGTTCGCTTACCCCGGCTGACCGGCTCATGAGCTCTGCTG 870
QY 481 CGGCTAGAGCTCTATGGCTGCTCTGGAGGATGAGCTCTGCTTACACGCGCCCTGTG 540
DB 871 CGGCTAGAGCTCTATGGCTGCTCTGGAGGATGAGCTCTGCTTACACGCGCCCTGTG 930
QY 541 GGGCAGACAATGATTTATCTGAGGCGCTGTACCTCAACGACTCCACCTATGACGAGCAT 600
DB 931 GGGCAGACAATGATTTATCTGAGGCGCTGTACCTCAACGACTCCACCTATGACGAGCAT 990
QY 601 ACCGTGGGCGGACTGAGTATGGGCTGTGGCCAGCTGGCAGATGCTGCTGGGCTG 660
DB 991 ACCGTGGGCGGACTGAGTATGGGCTGTGGCCAGCTGGCAGATGCTGCTGGGCTG 1050
QY 661 GATGACTTTAGGAAGAGTCAGGAGCTGGGCTGTGGCCAGCTGGCAGATGCTGCTGGGATG 720
DB 1051 GATGACTTTAGGAAGAGTCAGGAGCTGGGCTGTGGCCAGCTGGCAGATGCTGCTGGGATG 1110
QY 721 AGCAACACAGCTTCTCAGTGGCTATGTGGAGATGGAGTTGAGTTGACCGGCTGAGG 780
DB 1111 AGCAACACAGCTTCTCAGTGGCTATGTGGAGATGGAGTTGAGTTGACCGGCTGAGG 1170
QY 781 GCGTTCAGGCTATGAGTCCACTGTATACACATGACAGCTGGGAGCCGCTGTGCT 840
DB 1171 GCGTTCAGGCTATGAGTCCACTGTATACACATGACAGCTGGGAGCCGCTGTGCT 1230
QY 841 GCGGGGTGGAATGTCGCTTCGCGGCTGGCCCTGCATGGCTGGGAGGGGAGCCCATG 900
DB 1231 GCGGGGTGGAATGTCGCTTCGCGGCTGGCCCTGCATGGCTGGGAGGGGAGCCCATG 1290
QY 901 CGCCACAACCTAGGGGGAACCTCGGGGACCCAGAGCCGCGGCTGCTCAGTGCCTT 960
DB 1291 CGCCACAACCTAGGGGGAACCTCGGGGACCCAGAGCCGCGGCTGCTCAGTGCCTT 1350
QY 961 GCGGCGGTGGCTTCGCTTCTCAGTGGCGCTTCTTTCGCGGCGCCCTGCTTACT 1020
DB 1351 GCGGCGGTGGCTTCGCTTCTCAGTGGCGCTTCTTTCGCGGCGCCCTGCTTACT 1410
QY 1021 TTCAGGAATCTCTTCTATCTCTGATGTGGTGAACATTCCTCCCGGACTGGAGGC 1080
DB 1411 TTCAGGAATCTCTTCTATCTCTGATGTGGTGAACATTCCTCCCGGACTGGAGGC 1470
QY 1081 ACCTTCCCGGACGCGCTTGGTGGCGCTGGCCACCTCCACCACTTCAGCAGCTG 1140
DB 1471 ACCTTCCCGGACGCGCTTGGTGGCGCTGGCCACCTCCACCACTTCAGCAGCTG 1530

QY 1141 GAGCTGGAGCCAGAGCCAGCAGCCCGTGGCCAGCCGAGGGAGCCCGACCGCC 1197
DB 1531 GAGCTGGAGCCAGAGCCAGCAGCCCGTGGCCAGCCGAGGGAGCCCGACCGCC 1587
RESULT 5
HUMCAKA
LOCUS
DEFINITION Homo sapiens Cak receptor kinase mRNA, complete cds.
ACCESSION U57508
NID 91160924
VERSION 1.1
KEYWORDS 3692 bp mRNA
SOURCE Homo sapiens Cak receptor kinase mRNA, complete cds.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3692)
AUTHORS Perez, J.L., Jing, S.Q. and Wong, T.W.
TITLE Identification of two isoforms of the Cak receptor kinase that are
coexpressed in breast tumor cell lines
JOURNAL Oncogene 12 (7), 1469-1477 (1996)
MEDLINE 96204002
FEATURES
Location/Qualifiers
1..3692
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="lung"
/note="(vector lambda gt10)"
193..2823
/gene="Cak"
193..2823
/gene="Cak"
/codon_start=1
/product="receptor kinase"
/protein_id="AAB05208.1"
/db_xref="PID:g1160925"
/db_xref="GI:1160925"
/translation="MGPEALSSLLLLVASGDADMGHFDPAKRYALGMQDRTIPD
SDISASSWSDTAARHSRLSSDGDGAWCPAGSVFPKEEYLQVLDLRLHLVALVGT
QKHAGLGKGFESRISRLYSDRRMRWMDKRWQGEVIGSNGEDPEGVVLDLGGPMV
AKLRFYPRADRVMSVCLRLDGLSLTAPVGTQMYLSEAYLNDSTYDGH
TVGQGLGQLADGVLDLDFRKSQELRWPMGYDYVGSWNSFSSCYVMEFEFDR
LFAQAMQVHCNMHTLGLARLPQVECFRFRGPMAMWEGEMRNLGNLPPRARAV
SVPLGRVARELQCFELFAGPWLFLSEISFISDVVNSSPALGGTFPPAPWPPGPP
TNFSLLELPQCPQVAKAGSPFALLGLCLVAIILLLLILALMLRHLRLLSKA
ERRVLEELTVHLSVPGDTLLNNRPPEPPPYOEPNPNPHSPHSPVNGSAYS
DYMEPEKPGAPLLPPPPQNSVPHYAADIIVTQVGTGNTYAVPALPGVADGPPRV
DPPRSLRFKEKLGQFCEVHLCESVDSPQDLVSLDFPLNVRKHPHLLVAVKILPDA
TKNARNDLFKEVIMSRLKDPNIIRLLGVQDDPLCMITDYMENGLNQLSAHOLE
DKAAGAPCDGAAOGPTISYPMLLHVAQAISGNRYLATLNFVHRDLATNRLNGEN
FTIKTADGMSRLNLYAGIYRVQGRAVLPFRMAWECILMGKFTTASDVWAFVGLWE
VMLPCRAQFPGQLTDEVIENAGEFFRDGROVYLSRPPACPGQLYELMLRCWRESE
QRPFPFSLHRLAEALNTV"
BASE COUNT 704 a 1127 c 1102 g 759 t
ORIGIN
Query Match 99.9%; Score 1195.4; DB 10; Length 3692;
Best Local Similarity 99.9%; Pred. No. 1.5e-248;
Matches 1196; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GATGCTGACATGAAGGACATTTGATCTGCTGCAAGTGCCTGCGCATGCCATGCCATGCCAG 50
DB 247 GATGCTGACATGAAGGACATTTGATCTGCTGCAAGTGCCTGCGCATGCCATGCCAG 306
QY 61 GACCGGACCATCCAGACAGTACATCTGCTGCTCAGCTCCTGCTGCTGCTGCTGCTGCTG 120
DB 307 GACCGGACCATCCAGACAGTACATCTGCTGCTCAGCTCCTGCTGCTGCTGCTGCTGCTG 366
QY 121 GCGCGCCACAGCAGGTTGGAGAGCAGTACGCGGATGGGCGCTGCTGCTGCTGCTGCTGCTG 180
DB 367 GCGCGCCACAGCAGGTTGGAGAGCAGTACGCGGATGGGCGCTGCTGCTGCTGCTGCTGCTG 426

QY 181 GTGTTTCCAGAGGAGGAGTACTTGCAGGTGGATCTACAAGACTGCACCTGGTGGCT 240
DB 427 GTGTTTCCAGAGGAGGAGTACTTGCAGGTGGATCTACAAGACTGCACCTGGTGGCT 486
QY 241 CTGTTGGGACCCAGGAGGAGGAGTACTTGCAGGTGGATCTACAAGACTGCACCTGGTGGCT 300
DB 487 CTGTTGGGACCCAGGAGGAGGAGTACTTGCAGGTGGATCTACAAGACTGCACCTGGTGGCT 546
QY 301 CGGTGCTTACTCCCGGATGTCGCGCTGGATGGCTGGAAGAGCCCTGGGTGAG 360
DB 547 CGGTGCTTACTCCCGGATGTCGCGCTGGATGGCTGGAAGAGCCCTGGGTGAG 606
QY 361 GAGGTGATCTCAGGCAATGAGGAGCCCTGAGGAGTGTGCTGAAGAGCCCTGGGTGAG 420
DB 607 GAGGTGATCTCAGGCAATGAGGAGCCCTGAGGAGTGTGCTGAAGAGCCCTGGGTGAG 666
QY 421 ATGTTGCGCCGACTGCTGCTTACCCCGGCTGACCGGCTCATGAGGCTGCTGCTG 480
DB 667 ATGTTGCGCCGACTGCTGCTTACCCCGGCTGACCGGCTCATGAGGCTGCTGCTG 726
QY 481 CGGTGAGAGCTTATGCTGCTTGGAGGATGAGTCTGCTTACACCGCCCTGCTG 540
DB 727 CGGTGAGAGCTTATGCTGCTTGGAGGATGAGTCTGCTTACACCGCCCTGCTG 786
QY 541 GGGCAGACAATGATTAATCTAGAGCCGCTGACCTCAACGACTCCACCTATGAGGAGAT 600
DB 787 GGGCAGACAATGATTAATCTAGAGCCGCTGACCTCAACGACTCCACCTATGAGGAGAT 846
QY 601 ACCGTGGCGGACCTGAGTATGGGGCTGCGGCTGAGGATGCTGCTGAGGCTG 660
DB 847 ACCGTGGCGGACCTGAGTATGGGGCTGCGGCTGAGGATGCTGCTGAGGCTG 906
QY 661 GATGACTTTAGGAAGAGTACGAGCTGCGGCTGCGGCTGAGGATGCTGCTGAGGCTG 720
DB 907 GATGACTTTAGGAAGAGTACGAGCTGCGGCTGCGGCTGAGGATGCTGCTGAGGCTG 966
QY 721 AGCAACACAGCTTCTCAGTGGCTATGTGAGATGAGTGTGAGTTGACCGGCTGAGG 780
DB 967 AGCAACACAGCTTCTCAGTGGCTATGTGAGATGAGTGTGAGTTGACCGGCTGAGG 1026
QY 781 GCTTCCAGGCTATGCAAGTCCACTGTATCAACATGACAGCTGGAGCCGCTGCTGCT 840
DB 1027 GCTTCCAGGCTATGCAAGTCCACTGTATCAACATGACAGCTGGAGCCGCTGCTGCT 1086
QY 841 GCGGGGTGGAATGCTGCTTCCGCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 900
DB 1087 GCGGGGTGGAATGCTGCTTCCGCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 1146
QY 901 GCGGACACCTAGGGGCAACCTGCGGCAACCTGCGGCAACCTGCGGCTGCGGCTGCGGCTG 960
DB 1147 GCGGACACCTAGGGGCAACCTGCGGCAACCTGCGGCAACCTGCGGCTGCGGCTGCGGCTG 1206
QY 961 GCGGCGCTGCTGCTTCTGAGTGGCTGCTGCTTCTGAGTGGGCTGCTGCTGCTGCTGCT 1020
DB 1207 GCGGCGCTGCTGCTTCTGAGTGGCTGCTGCTTCTGAGTGGGCTGCTGCTGCTGCTGCT 1266
QY 1021 TTCACGAAATCTCCTTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
DB 1267 TTCACGAAATCTCCTTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1326
QY 1081 ACCTTCCGCGCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
DB 1327 ACCTTCCGCGCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1386
QY 1141 GAGCTGGAGCCAGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1197
DB 1387 GAGCTGGAGCCAGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1443

RESULT 6
I68738 168738 3962 bp DNA PAT 23-DEC-1997
LOCUS

DEFINITION Sequence 1 from patent US 5677144.
ACCESSION I68738
NID 92830860
VERSION I68738.1 GI:2830860
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3962)
AUTHORS Ullrich, A. and Alves, F. Hildegard Elisabeth.
TITLE Recombinant DNA encoding CK2, a receptor tyrosine kinase
JOURNAL Patent: US 5677144-A 1 14-Oct-1997;
FEATURES Location/Qualifiers
source 1..3962
BASE COUNT 735 a 1234 c 1182 g 811 t
ORIGIN

Query Match 99.6%; Score 1192.2; DB 5; Length 3962;
Best Local Similarity 99.7%; Pred. No. 7.2e-248;
Matches 1194; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 GATGCTGACATCAAGGAGCATTTTGCATCTGCCAAGTCCCGCTATGCCCTGGGCATGCAG 60
DB 375 GATGCTGACATCAAGGAGCATTTTGCATCTGCCAAGTCCCGCTATGCCCTGGGCATGCAG 434
QY 61 GACCGGACCATCCAGACAGTGCATCTCTGCTTCCAGCTCTCTGCTCAGATTCCACTGCC 120
DB 435 GACCGGACCATCCAGACAGTGCATCTCTGCTTCCAGCTCTCTGCTCAGATTCCACTGCC 494
QY 121 GCGCGCCACAGCAGGTTGGAGAGCAGTGCAGGGGATGGGGCTTGGTGCCTGGCGAGGGTGC 180
DB 495 GCGCGCCACAGCAGGTTGGAGAGCAGTGCAGGGGATGGGGCTTGGTGCCTGGCGAGGGTGC 554
QY 181 GTGTTTCCCAAGAGGAGGAGTACTTGCAGTGGATCTACAACGACTGCACCTGGTGGCT 240
DB 555 GTGTTTCCCAAGAGGAGGAGTACTTGCAGTGGATCTACAACGACTGCACCTGGTGGCT 614
QY 241 CTGTTGGGCAACCCAGGAGCGCATGCCGGGGCTTGGGCAAGAGTCTTCCCGGAGCTAC 300
DB 615 CTGTTGGGCAACCCAGGAGCGCATGCCGGGGCTTGGGCAAGAGTCTTCCCGGAGCTAC 674
QY 301 CGGCTGCTTACTCCCGGATGGTGCCTGCTGGATGGCTGGAAGAGCCGCTGGGGTGCAG 360
DB 675 CGGCTGCTTACTCCCGGATGGTGCCTGCTGGATGGCTGGAAGAGCCGCTGGGGTGCAG 734
QY 361 GAGGTGATCTCAGGCAATGAGGAGCCCTGAGGAGTGGTGTGAAGAGCTTGGGCCCCC 420
DB 735 GAGGTGATCTCAGGCAATGAGGAGCCCTGAGGAGTGGTGTGAAGAGCTTGGGCCCCC 794
QY 421 ATGTTTCCCGGACTGGTTCGCTTCTACCCCGGCTGACCGGCTCATGAGGCTCTGTCTG 480
DB 795 ATGTTTCCCGGACTGGTTCGCTTCTACCCCGGCTGACCGGCTCATGAGTCTGTCTG 854
QY 481 CGGCTGAGAGCTTATGCTGCTGCTGCTGAGGAGTGAAGTCTGCTTACACCGCCCTGTG 540
DB 855 CGGCTGAGAGCTTATGCTGCTGCTGCTGAGGAGTGAAGTCTGCTTACACCGCCCTGTG 914
QY 541 GGGCAGACAATGATTAATCTGAGGCGCTGACCTCAACGACTCCACCTATGACGAGAT 600
DB 915 GGGCAGACAATGATTAATCTGAGGCGCTGACCTCAACGACTCCACCTATGACGAGAT 974
QY 601 ACCGTGGCGGACTGCAGTATGGGGCTTGGGCCAGCTGGCAGATGGTGTGGTGGGGTGC 660
DB 975 ACCGTGGCGGACTGCAGTATGGGGCTTGGGCCAGCTGGCAGATGGTGTGGTGGGGTGC 1034
QY 661 GATGACTTTAGGAAGAGTGCAGGAGTGGGGTCTGGCCAGGCTATGCTATGTTGGGATGG 720
DB 1035 GATGACTTTAGGAAGAGTGCAGGAGTGGGGTCTGGCCAGGCTATGCTATGTTGGGATGG 1094
QY 721 AGCAACACAGCTTCTCAGTGGCTATGAGAGTGGAGTTTGGTGGAGCTGAGGCTGAGG 780

```

Db 1095 AGCAACCACAGCTTCTCCAGTGGCTATGTGGAGATGGAGTTTGAGTTTGACCGGCTGAGG 1154
Qy 781 GCCTTCAGGCTATGACAGGTCACCTGTAACAAATGACACACGCTGGAGCCGCTGTCCT 840
Db 1155 GCCTTCAGGCTATGACAGGTCACCTGTAACAAATGACACACGCTGGAGCCGCTGTCCT 1214
Qy 841 GCGGGGTGAATGTCCTTCGCGGTGGCCCTGCCATGCCCTGGAGGGGAGCCCATG 900
Db 1215 GCGGGGTGAATGTCCTTCGCGGTGGCCCTGCCATGCCCTGGAGGGGAGCCCATG 1274
Qy 901 CGCCACAACCTAGGGGCAACCTGGGGGAGCCGAGAGCCGCGGCTCTCAGTGCCCTT 960
Db 1275 CGCCACAACCTAGGGGCAACCTGGGGGAGCCGAGAGCCGCGGCTCTCAGTGCCCTT 1334
Qy 961 GCGGGCGTGTGGCTCGCTTTCGAGTGGCGCTTCTCTTTCCGGGGCCCTGTTTACTC 1020
Db 1335 GCGGGCGTGTGGCTCGCTTTCGAGTGGCGCTTCTCTTTCCGGGGCCCTGTTTACTC 1394
Qy 1021 TTACGCGAATCTCTTCATCTCTGATGGTGAACAAATTCCTCTCCGGCACTGGAGGC 1080
Db 1395 TTACGCGAATCTCTTCATCTCTGATGGTGAACAAATTCCTCTCCGGCACTGGAGGC 1454
Qy 1081 ACCTTCGCGCAGCCCTCGTGGCGCCCTGGCCGACCTCCACCAACTTCAGCAGCTTG 1140
Db 1455 ACCTTCGCGCAGCCCTCGTGGCGCCCTGGCCGACCTCCACCAACTTCAGCAGCTTG 1514
Qy 1141 GAGGTGAGCCAGAGGCCAGCAGCCGCTGGCCAAAGCCGAGGGAGCCGCGCC 1197
Db 1515 GAGGTGAGCCAGAGGCCAGCAGCCGCTGGCCAAAGCCGAGGGAGCCGCGCC 1571

RESULT 7
HSTRKE HSTRKE 3554 bp mRNA PRI 23-JAN-1997
LOCUS H. sapiens TRK E mRNA.
DEFINITION X74979
ACCESSION X74979
NID 9400462
VERSION X74979.1 GI:400462
KEYWORDS trk E gene; Trk E protein.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3554)
AUTHORS Di Marco, E., Cutuli, N., Guerra, L., Cancedda, R. and De Luca, M.
TITLE Molecular cloning of trkE, a novel trk-related putative tyrosine
kinase receptor isolated from normal human keratinocytes and widely
expressed by normal human tissues
J. Biol. Chem. 268 (32), 24290-24295 (1993)
JOURNAL 9403265
MEDLINE 2 (bases 1 to 3554)
REFERENCE De Luca, M.
AUTHORS Direct Submission
TITLE Submitted (08-SEP-1993) M. De Luca, Istituto Nazionale per la
JOURNAL Ricerca sul Cancro, Lab Differenziamento Cellulare, Viale
Benedetto XV, 10 - 16132 Genova, ITALY
Location/Qualifiers
1. .3554
/organism="Homo sapiens"
/db_xref="taxon:9606"
/dev_stage="germ line, fetal /adult"
/tissue_type="brain/keratinocytes"
46. .2676
/gene="trke"
/codon_start=1
/product="trke"
/protein_id="CAA52915.1"
/db_xref="PID:9400463"
/db_xref="GI:400463"
/db_xref="SWISS-PROT:008345"
/translation="MGPEALSSLLLLLVASGDADMKGHFDPKCRVALGMDRTIPD

```

```

SDIASSSWSDSTAARSHRLESDDGAMCPAGSVFPEKEEYLQVDLQRLHLVALVGT
QGRHAGLGKEFSRSELYSRDGRWMMGHKDRWQEVISGNEDEPGYVLKDLGPPMV
ARLYRYPADRVMSVCLRVLYGCLWRDGLSTYAPVGTMYLSSEAYLNDLSDGH
TVGGLQGGGLQGLADGVLDLDFKSOELRWPGYDYVGNWSHFSYSGYVEMFEFDR
LRAFQAMQVHCNNHMTLGLARLPGVCEFRFRGPMAMWEGEPHNNLGGNLDPRARV
SVPLGGRVAREFLQCFRFLFAGPWLIFSEIFSDVWNNSSPALGCTPPAPWPPGPP
TNFSSLELEPRGQOPVAKAGSPTAILIGCLVALILLLLLI IALMLRHLWRLLSKA
ERRVLEBELTVHLSVPDGTILINRPGPREPPYQPRPGNPPHSPCPVNGSAYSG
DYMEPEKGPAPLLPPPPQNSVPHYAEADIVTLOGVTGGNTYAPVAPPGVGGPPRV
DPRSLRFLKEKLGEGFGEVHUCEVDSPDLVSLDFPLNVRKHGHLNVAVKTLRPPDA
TKNARNDLFKEKIMSKRLDPNIRLLGVCVDDPLCMITDYMENGLDGLVSAHOLE
DKAAGAPGQDGAQOQFTISYPMLLHVAQIASGMRYLATLNFVETDLATRNCLVGEN
FTIKIADFGMSRLYAGDYRVQGRVLPIRWMAWECILMGKFTTASDVYAFQVTLWE
VLMCRAOPFGQLTDQVIENAGEFFRDGGRQVYLSRPPACPOGLYELMLRCSRESE
ORPPESQLHREFAEDALNTV"
BASE COUNT 682 a 1064 c 1065 g 743 t
ORIGIN
Query Match 99.6%; Score 1192.2; DB 10; Length 3554;
Best Local Similarity 99.7%; Pred. No. 7.3e-248;
Matches 1194; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 GATGCTGACATGAAGGGACATTTTGTCTCTGCCAAGTCCGCTATGCCCTGGGCATGCAG 60
Db 100 GATGCTGACATGAAGGGACATTTTGTCTCTGCCAAGTCCGCTATGCCCTGGGCATGCAG 159
Qy 61 GACCGGACCATCCACAGACAGTACATCTGTCTTCCAGCTCCTGGTTCAGATTCACATGCC 120
Db 160 GACCGGACCATCCACAGACAGTACATCTGTCTTCCAGCTCCTGGTTCAGATTCACATGCC 219
Qy 121 GCCCGCCACAGCAGTGGAGAGCAGTACGCGGATGGGCGCTGGTCCCGCAGGAGTCG 180
Db 220 GCCCGCCACAGCAGTGGAGAGCAGTACGCGGATGGGCGCTGGTCCCGCAGGAGTCG 279
Qy 181 GTGTTTCCCAAGGAGGAGTACTTTCAGTGGATCTACAACGACTSCACCTGGTGGCT 240
Db 280 GTGTTTCCCAAGGAGGAGTACTTTCAGTGGATCTACAACGACTSCACCTGGTGGCT 339
Qy 241 CTGCTGGGCCACAGGAGGAGCAGTCCGCGGCGCTGGGCAAGAGTTCCTCCCGAGGATAC 300
Db 340 CTGCTGGGCCACAGGAGGAGCAGTCCGCGGCGCTGGGCAAGAGTTCCTCCCGAGGATAC 399
Qy 301 CGGCTGGGTACTCCCGGATGGTCCGCTGGATGGCTGGAAGGACCCCTGGGGTCAG 360
Db 400 CGGCTGGGTACTCCCGGATGGTCCGCTGGATGGCTGGAAGGACCCCTGGGGTCAG 459
Qy 361 GAGTGATCTCAGCAATAGGACCCCTGAGGAGTGTGTCTGAAGGACCTTGGGCCCCCC 420
Db 460 GAGTGATCTCAGCAATAGGACCCCTGAGGAGTGTGTCTGAAGGACCTTGGGCCCCCC 519
Qy 421 ATGTTTCCCGACATGTTCTCTTACCCCGGCTGACCGGCTCATGACGGTCTGTCTG 480
Db 520 ATGTTTCCCGACATGTTCTCTTACCCCGGCTGACCGGCTCATGACGGTCTGTCTG 579
Qy 481 CGGCTAGAGCTCTATGGCTGCCTCTGGAGGATGGACTCTCTTACACCGCCCTGTG 540
Db 580 CGGCTAGAGCTCTATGGCTGCCTCTGGAGGATGGACTCTCTTACACCGCCCTGTG 639
Qy 541 GGGCAGACATGATTTATCTGAGGCCGTGTACTCTAACGACTCCACCTATGACGGACAT 600
Db 640 GGGCAGACATGATTTATCTGAGGCCGTGTACTCTAACGACTCCACCTATGACGGACAT 699
Qy 601 ACCGTGGCGGACATGACAGTATGGGGTCTGGGCCAGCTGGCAGATGGTGTGGGGCTG 660
Db 700 ACCGTGGCGGACATGACAGTATGGGGTCTGGGCCAGCTGGCAGATGGTGTGGGGCTG 759
Qy 661 GATGACTTTAGGAAGAGTACAGGAGCTCGCGGTCTGGCCAGGCTATGACTATGTGGATGG 720
Db 760 GATGACTTTAGGAAGAGTACAGGAGCTCGCGGTCTGGCCAGGCTATGACTATGTGGATGG 819
Qy 721 AGCAACACAGCTTCTCCAGTGGCTATGTGGAGATGGAGTTTGAGTTTGACCGGCTGAGG 780

```



```

Db 820 AGCAACACAGCTTCTCCAGTGGCTATGTGGAGATGGAGTTGAGTTCGACCCGCGTGGAG 879
Qy 781 GCCTTCCAGGCTATGAGTGCACATGTAACACATGCACAGCTGGAGCCGCTGCTGCT 840
Db 880 GCCTTCCAGGCTATGAGTGCACATGTAACACATGCACAGCTGGAGCCGCTGCTGCT 939
Qy 841 GCGGGGTGGAATGTCGTTCCGGCGTGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db 940 GCGGGGTGGAATGTCGTTCCGGCGTGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 999
Qy 901 GCGCACACCTAGGGGCAACCTGGGGACCCAGAGCCCGGCTGCTGCTGCTGCTGCTGCTGCT 960
Db 1000 GCGCACACCTAGGGGCAACCTGGGGACCCAGAGCCCGGCTGCTGCTGCTGCTGCTGCTGCT 1059
Qy 961 GCGGGCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
Db 1060 GCGGGCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1119
Qy 1021 TTACAGGAATCTCTTCTATCTCTGATGTGGTGAACAATTCCTTCCGGCACTGGGAGGC 1080
Db 1120 TTACAGGAATCTCTTCTATCTCTGATGTGGTGAACAATTCCTTCCGGCACTGGGAGGC 1179
Qy 1081 ACCTTCCCGCCAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
Db 1180 ACCTTCCCGCCAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1239
Qy 1141 GAGCTGGAGCCAGAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1197
Db 1240 GAGCTGGAGCCAGAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1296

RESULT 8
HUMRTK HUMRTK 3738 bp mRNA PRI 27-SEP-1993
LOCUS Homo sapiens receptor tyrosine kinase mRNA, complete cds.
DEFINITION L11315
ACCESSION 9403386
NID L11315.1
VERSION 1
KEYWORDS receptor tyrosine kinase.
SOURCE Homo sapiens (library: lambda-gt11) term placenta cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE Johnson, J.D., Edman, J.C. and Rutter, W.J.
AUTHORS A receptor tyrosine kinase found in breast carcinoma cells has an
TITLE extracellular discoidin I-like domain
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 90, 5677-5681 (1993)
MEDLINE 93296201
COMMENT On Sep 27, 1993 this sequence version replaced gi:307397.
FEATURES
Location/Qualifiers
source
1..3738
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="term placenta"
/tissue_lib="lambda-gt11"
1..141
142..2883
/standard_name="discoidin I-like domain receptor tyrosine
kinase"
/codon_start=1
/product="receptor tyrosine kinase"
/protein_id="AAA02866.1"
/db_xref="PID:403387"
/db_xref="GI:403387"
/translation="MGPEALSSLLLLVAGSDADMKGHDPKCRVALGMDRTIPD
SDIASSSWSSTAAHRSRLSSDGDGAWCPAGSVFPEEYELQVDLQVRLHYALVGT
QGRHAGGUGKEFSRSYRLRYSRDGRWGMKDRMGQEVISGNDEPGEVYLKDLGPPWV
ARLVRFYPRADRVNVCYLRVLYGLWRDLGLSLTAPYQVNTMYSEAVYINDSYDGH
TVGGLQYGLGLGADVGLDQDFRKSQELRVWPGDYVVGWSNHSFSSGYVEMEFDFR
LRAFOAMVHCNMLTGLARPLFGVECFRRGPAMAWEGEPHNLGNLGDPRARV
SVPLGRVARELQCRFLFAGPWLFLSEISFSDVNVNNSPALGTFEPAPWPPGPP
TNFSSLELEPRQOPVAKAGSPTAILGCLVAIILLLLLLIALMLRHLWRLLSKA

```

5' UTR
CDS

```

ERRVLEELTVHLSVPGDTILNNRPGPREPPPYQEPREPRGNPNPHSPAPCVNPSALLL
SNPARYLLLATYARPPRGPPPTPAWAKPTNTQAYSGDYMEPKPAPLPPPPONSV
PHYAADLITVQGVGTGYAYPALPGVAGDPPRVDFPSRLRFLKEKLGFGFGEV
HICEYDSDQLVSLDFPLNVRKHGHLVAVILRPDATTNARNDFLKEVKIWSRLKDP
NIIRLLGVCVDDPLDMITDYMGNDLNOFLSAHOLEDKAAGAPGQQAQPTISY
PMLLHVAQIASGRMYLATLNFVHRDLATRNCLVGENFTIKIADFGSMRNLXAGDYR
VOGRAVLPIRNMWECILMGKFTTASDWAFGVIVWVLMCLRAQPGFQGLQDEQVIE
AGEFFRDQGRQVYLSRPPACPGQYELMLRCWSRESEQRPPFSQLHRFLAEDALNTV
2887..3738
3' UTR
polyA_site 3738
BASE COUNT 698 a 1140 c 1123 g 777 t
ORIGIN

```

```

Query Match 99.5%; Score 1190.6; DB 9; Length 3738;
Best Local Similarity 99.7%; Pred. No. 1.6e-247;
Matches 1193; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GATGCTGACATGAAGGACATTTTGTATCTGCCAAGTGCCTATGCTCCCTGGCATGCAG 60
Db 196 GATGCTGACATGAAGGACATTTTGTATCTGCCAAGTGCCTATGCTCCCTGGCATGCAG 255
Qy 61 GACCGGACCATCCACAGACAGTACATCTCTGCTTCCAGCTCCTGGTCAAGATCCCACTGCC 120
Db 256 GACCGGACCATCCACAGACAGTACATCTCTGCTTCCAGCTCCTGGTCAAGATCCCACTGCC 315
Qy 121 GCGCGCCACACAGAGTGGAGAGCAGTACGAGGAGTGGGGGCTGGTGGCCCGCAGGGTGC 180
Db 316 GCGCGCCACACAGAGTGGAGAGCAGTACGAGGAGTGGGGGCTGGTGGCCCGCAGGGTGC 375
Qy 181 GTGTTTCCCAAGGAGGAGTACTTTCAGGTGGATCTTCAACGAGTGCACCTGGTGGCT 240
Db 376 GTGTTTCCCAAGGAGGAGTACTTTCAGGTGGATCTTCAACGAGTGCACCTGGTGGCT 435
Qy 241 CTGTTGGGACCCACAGGACGCGCATGCCGGGGCTGGGAAGGAGTTCCTCCGGAGCTAC 300
Db 436 CTGTTGGGACCCACAGGACGCGCATGCCGGGGCTGGGAAGGAGTTCCTCCGGAGCTAC 495
Qy 301 CGGTCGCTTACTCCCGGATGTCGCCCGCTGGATGGGTGGAGGACCGCTGGGGTTCAG 360
Db 496 CGGTCGCTTACTCCCGGATGTCGCCCGCTGGATGGGTGGAGGACCGCTGGGGTTCAG 555
Qy 361 GAGTGTATCTCAGCAATGAGGACCTGAGGAGTGGTGTGCTGAAGGACCTTGGGCCCCC 420
Db 556 GAGTGTATCTCAGCAATGAGGACCTGAGGAGTGGTGTGCTGAAGGACCTTGGGCCCCC 615
Qy 421 ATGTTTCCCGACTGTTTCGCTTCTACCCCGGCTGACCGGCTCATGAGCGTCTGCTG 480
Db 616 ATGTTTCCCGACTGTTTCGCTTCTACCCCGGCTGACCGGCTCATGAGCGTCTGCTG 675
Qy 481 CGGTTAGAGTCTATGCTGCTGCTGAGGAGTGGACTCCTGCTTACACCGCCCTGTG 540
Db 676 CGGTTAGAGTCTATGCTGCTGCTGAGGAGTGGACTCCTGCTTACACCGCCCTGTG 735
Qy 541 GGGCAGACATGTTATTTATCTGAGGCGCTTACTCAACGACTCCACCTATGACGACAT 600
Db 736 GGGCAGACATGTTATTTATCTGAGGCGCTTACTCAACGACTCCACCTATGACGACAT 795
Qy 601 ACCGTGGCGGACTGCAATGATGGGGTCTGGCCAGCTGGCAGATGGTGTGGGGGCTG 660
Db 796 ACCGTGGCGGACTGCAATGATGGGGTCTGGCCAGCTGGCAGATGGTGTGGGGGCTG 855
Qy 661 GATGACATTTAGGAAGAGTCAAGGAGTGGGGTCTGGCCAGGCTATGACTATGTGGGATGG 720
Db 856 GATGACATTTAGGAAGAGTCAAGGAGTGGGGTCTGGCCAGGCTATGACTATGTGGGATGG 915
Qy 721 AGCAACACAGCTTCTCCAGTGGCTATGTGAGATGGATGAGTTTGACCGGCTGAGG 780
Db 916 AGCAACACAGCTTCTCCAGTGGCTATGTGAGATGGATGAGTTTGACCGGCTGAGG 975
Qy 781 GCGTTCACAGGCTATGCAAGGTCCTGTAACAACATGCACACGCTGGGAGCCGCTGCTGCT 840
Db 976 GCGTTCACAGGCTATGCAAGGTCCTGTAACAACATGCACACGCTGGGAGCCGCTGCTGCT 1035

```



```
QY 841 GCGGGGTGAATGTCCTTCGGCGCTGGCCCTGCCATGGCTGGAGGGGAGGCCCATG 900
Db 1036 GCGGGGTGAATGTCCTTCGGCGCTGGCCCTGCCATGGCTGGAGGGGAGGCCCATG 1095
QY 901 CGCCACAACCTAGGGGGCAACCTGGGGAGCCAGAGCCCGGGCTGTCTCACTGCCCTT 960
Db 1096 CGCCACAACCTAGGGGGCAACCTGGGGAGCCAGAGCCCGGGCTGTCTCACTGCCCTT 1155
QY 961 GCGGGCGTGTGCTCGCTTCCTTCAGTGCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 1020
Db 1156 GCGGGCGTGTGCTCGCTTCCTTCAGTGCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 1215
QY 1021 TTGAGGGAATCTCTTCATCTCTGATGTTGTAACAATTCCTTCCTTCCTTCCTTCCT 1080
Db 1216 TTGAGGGAATCTCTTCATCTCTGATGTTGTAACAATTCCTTCCTTCCTTCCTTCCT 1275
QY 1081 ACCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCT 1140
Db 1276 ACCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCT 1335
QY 1141 GAGCTGAGCCAGAGCCAGAGCCAGAGCCAGAGCCAGAGCCAGAGCCAGAGCCAGAG 1197
Db 1336 GAGCTGAGCCAGAGCCAGAGCCAGAGCCAGAGCCAGAGCCAGAGCCAGAGCCAGAG 1392

RESULT 9
A42378 3754 bp DNA PAT 06-MAR-1997
LOCUS
DEFINITION Sequence 1 from Patent WO9502187.
ACCESSION A42378
NID 92297854
VERSION A42378.1 GI:2297854
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Gusterson,B.A., Crompton,M.R., Mitchell,P.J., Barker,K.T.,
Martindale,J.E., Page,M.J. and Spence,P.
TITLE CELL GROWTH FACTOR RECEPTORS
JOURNAL Patent: WO 9502187-A 1 19-JAN-1995;
CANCER RES INST (GB)
COMMENT Other publication AU 7081094 950206.
FEATURES
Location/Qualifiers
1..3754
/organism="unidentified"
/db_xref="taxon:32644"
142..2886
/note="unnamed protein product"
/codon_start=1
/protein_id="CAA02688.1"
/db_xref="pid:e305343"
/db_xref="pid:g2297855"
/db_xref="gi:2297855"
/translation="MGPEALSSLLLLLVASGDADMKHDFDPKCRYALGMQDRTPD
SDISASSWSSTAAHRSLESDDGAWCPAGSVFPKEBEYLQVDLQRLHVALVGT
QGRHAGLGLKFSRSLRYSRDRGRWMDKRWQOEVI SGNEDPEGVVKLDLGPWV
ARLVFPYPRADRVMSVCLRLVGLYGLARDGLLSVTA PVGOTMYLSEAVYLNDSYDGH
TVGGVGLGLQADGVVGLDFRKSQELRWFPQYDXYVGNNSHSFSSGYVEMEFEDR
LRAQAMQVHCNNHHTLGLRPGVECFRFGPAMWEGEPHMRNLGNLGDPRARV
SVPLGVRVARELQCRFLFAGPWLFSFISDVYNNSSPALGTFPPAPWPPPPP
TNFSLLEPRGPQPPVAKAEGSPAILIGLVAIILLLIILALMLRWLRLLSK
AERLVBEELVHLSVPQDTILI NRPQPPPPYQPRPRGNPHSPACVPNGSALL
LSNPAYRLLIATYARPPRGPPTPAWAKPTQAYSGDYNEPEKFPALPPPPONS
VPHYAADIIVTLOGTYGNTYVAPLPGVGGPVPDFPRSLRFLKELGSGGFE
VHLEVDSPQDLSDPFLNVRKHPILLVAVILRPDATKNRDNFLKVKIMSLRK
PNIIRLVGVQDDPLCMITDYMENGLNQLSAHOLEDKAAEGACDGOAAGQPTIS
YPLLHVAQAQASGRMYRLATFNFLRDLATRNCLVGNFTIKIADFGMRLNLAGDYI
RVQGRVLPPIRWMAWECLMGKFTTASDVWAFVTLWEVLMLCRAOPFGOLTDEQVIE
NAGEFFRDQGRQVYLSRPPACPOGLYELMLRCNSRESESRPPFSQLRFLAEDALNTV
"

BASE COUNT 712 a 1146 c 1121 g 775 t
```

```
ORIGIN
Query Match 98.5%; Score 1179.2; DB 5; Length 3754;
Best Local Similarity 99.5%; Pred. No. 4.6e-245;
Matches 1194; Conservative 0; Mismatches 3; Indels 3; Gaps 1;

QY 1 GATGCTCACATGAAGGACATTTTGATCTGCAAGTGGCGCTATGCCCTGGGCATGCAG 60
Db 196 GATGCTCACATGAAGGACATTTTGATCTGCAAGTGGCGCTATGCCCTGGGCATGCAG 255
QY 61 GACCGGACCATCCAGACAGTACATCTCTGCTTCCAGCTCCTCTGGTCAGATTCACCTGCC 120
Db 256 GACCGGACCATCCAGACAGTACATCTCTGCTTCCAGCTCCTCTGGTCAGATTCACCTGCC 315
QY 121 GCGCGCCACAGCAGGTGGAGAGCAGTACGAGGGGATGGGCGCTGGTGGCCCGCAGGTCG 180
Db 316 GCGCGCCACAGCAGGTGGAGAGCAGTACGAGGGGATGGGCGCTGGTGGCCCGCAGGTCG 375
QY 181 GTGTTTCCCAAGGAGGAGGAGTACTTGCAGGTGATCTACAACGACTGCACCTGGTGGCT 240
Db 376 GTGTTTCCCAAGGAGGAGGAGTACTTGCAGGTGATCTACAACGACTGCACCTGGTGGCT 435
QY 241 CTGCTGGGCAACCAGGACGGCATGCCGGGGCTTGGCAAGAGTCTCTCCCGAGCTAC 300
Db 436 CTGCTGGGCAACCAGGACGGCATGCCGGGGCTTGGCAAGAGTCTCTCCCGAGCTAC 495
QY 301 CGGCTGGCTTACTCCCGGGATGTCGCGCTGGATGGCTGGAGGACCGCTGGGTCAG 360
Db 496 CGGCTGGCTTACTCCCGGGATGTCGCGCTGGATGGCTGGAGGACCGCTGGGTCAG 555
QY 361 GAGGTGATCTCAGGCAATGAGGAGCCTGAGGGAGTGGTCTGAAGGACCTTGGGCCCCC 420
Db 556 GAGGTGATCTCAGGCAATGAGGAGCCTGAGGGAGTGGTCTGAAGGACCTTGGGCCCCC 615
QY 421 ATGTTGCCCGACTGCTTCCTTACCCCGGGCTGACCGGCTCATGAGCGTCTCTCTG 480
Db 616 ATGTTGCCCGACTGCTTCCTTACCCCGGGCTGACCGGCTCATGAGCGTCTCTCTG 675
QY 481 CGGCTAGAGCTCTATGGCTGCTCTGGAGGATGGACTCTGTCTTACACCGCCCTGTG 540
Db 676 CGGCTAGAGCTCTATGGCTGCTCTGGAGGATGGACTCTGTCTTACACCGCCCTGTG 735
QY 541 GGCAGACAAATGATTTATCTGAGCGCTGTACTCTCAACGACTCCACCTATGACGACAT 600
Db 736 GGCAGACAAATGATTTATCTGAGCGCTGTACTCTCAACGACTCCACCTATGACGACAT 795
QY 601 ACCGTGGCGGACTGCAGTATGGGGTCTGGGCCAGCTGGCAGATGCTGTGGTGGGCTG 660
Db 796 ACCGTGGCGGACTGCAGTATGGGGTCTGGGCCAGCTGGCAGATGCTGTGGTGGGCTG 855
QY 661 GATGACTTTAGGAGAGTACAGGAGCTGGCGGCTTGGCCAGGCTATGACTATGTGGATGG 720
Db 856 GATGACTTTAGGAGAGTACAGGAGCTGGCGGCTTGGCCAGGCTATGACTATGTGGATGG 915
QY 721 AGCAACACAGCTTCTCCAGTGGCTATGTGGAGATGAGTGTGAGTTTACCGGCTGAGG 780
Db 916 AGCAACACAGCTTCTCCAGTGGCTATGTGGAGATGAGTGTGAGTTTACCGGCTGAGG 975
QY 781 GCCTTCCAGGCTATGACAGTTCCTACTGTAACAACATGCACACGCTGGAGCCCGCTGCCT 840
Db 976 GCCTTCCAGGCTATGACAGTTCCTACTGTAACAACATGCACACGCTGGAGCCCGCTGCCT 1035
QY 841 GCGGGGTGAATGTCCTTCGGCGCTGGCCCTGCCATGGCTGGAGGGGAGGCCCATG 900
Db 1036 GCGGGGTGAATGTCCTTCGGCGCTGGCCCTGCCATGGCTGGAGGGGAGGCCCATG 1095
QY 901 CGCCACAACCTAGGGGGCAACCTGGGGAGCCAGAGCCCGGGCTGTCTCACTGCCCTT 960
Db 1096 CGCCACAACCTAGGGGGCAACCTGGGGAGCCAGAGCCCGGGCTGTCTCACTGCCCTT 1155
QY 961 GCGGGCGTGTGCTCGCTTCCTTCAGTGCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 1020
```

```
Db 1156 GCGCGCGTGTGGCTCCTCTTCTGACGTGCGCGCTTCTCTTTCGGGGGCGCTGTTACTC 1215
Qy 1021 TTACGCAAAATCTCTTCACTCTGATGTGGTGAACAAATCTCTCCGGCACTGGGAGC 1080
Db 1216 TTACGCAAAATCTCTTCACTCTGATGTGGTGAACAAATCTCTCCGGCACTGGGAGC 1275
Qy 1081 ACCTTCCGGCAGCCCCCTGTGGTGGCCCTGGCCACCTCCCAACCACTTACAGACTTG 1140
Db 1276 ACCTTCCGGCAGCCCCCTGTGGTGGCCCTGGCCACCTCCCAACCACTTACAGACTTG 1335
Qy 1141 GAGCTGAGCCAGAGCCAGC---ACCCGCTGGCCAAAGCCCGAGGGAGCCGACCGCC 1197
Db 1336 GAGCTGAGCCAGAGCCAGCCAAAGCCCGTGGCCAAAGCCCGAGGGAGCCGACCGCC 1395

RESULT 10
AF026259 3659 bp mRNA ROD 27-OCT-1997
LOCUS Mus musculus receptor-like tyrosine kinase (Nep) mRNA, complete
DEFINITION cds.
ACCESSION AF026259
NID 92564939
VERSION AF026259.1 GI:2564939
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 3659)
Zerlin,M., Julius,M.A. and Goldfarb,M.
NEP: a novel receptor-like tyrosine kinase expressed in
proliferating neuroepithelia
Oncogene 8 (10), 2731-2739 (1993)
93390947
MEDLINE
2 (bases 1 to 3659)
Zerlin,M., Julius,M.A. and Goldfarb,M.
Direct Submission
Submitted (22-SEP-1997) Brookdale Center for Developmental and
Molecular Biology, Mount Sinai School of Medicine, 1 Gustave Levy
Place, New York, NY 10029, USA
Location/Qualifiers
FEATURES
source
1..3659
/organism="Mus musculus"
/strain="129"
/db_xref="taxon:10090"
/cell_type="CCE embryonic stem cells,
differentiation-induced"
1..3659
/gene="Nep"
/101..2836
/gene="Nep"
/feature="NEP; DDRL"
/codon_start=1
/product="receptor-like tyrosine kinase"
/protein_id="A898186.1"
/db_xref="PIR:92564940"
/db_xref="GI:2564940"
/translation="MGTGLSLLLLLVITGDAMKGFPAKCYALQMQRDRTIP
DSQTSVSSWSSTAAHRSLESDDGAWCPAGVPFKEEYLVQDLRLRLHLVALVG
TOGRHAGLGCKEFSRSLRLAYSRDRRMWDKRWGQEVISGNEDPGVYVLDLPGPM
VARLRYPRADRVMSVCLVELYGLWRDGLLSYAPVQTMQLSEVAVHLNDSYD
GYTAGQYGLQGLADGVVGLDDFRQSELVRWPGIDYVGSNSQSFPTGYVEMFEF
DRLTFTMGVHNNMHTLGLRFGVECFKRGPAWEGVEPRHALGSLGDPRAR
AISVPLGHVGRFLQCFRFLFAGPLPSEISFISDVNDSSTFPFPAWPPGPPPTN
FSSLEPRQGPVGRAGESPAILIGLVAIIILLIIALLMLRLHWRLLSADS
RVLEELTVHLSVPGDPTILNNRGPPEPPYOEPRPGTPPHSAPCVNPNGALLSN
PARYLLIATVAPRPGPTPAWKPNTQACSGDYMEEKPGAPLLPPQNSVPH
YAERDILVATGGNTIYAPALPPGAVGDPVDFPRSLRFRKLGEGQGEVHL
CEVEDPDQLVSSDFPISVHKGHLLVAVKILRPDANKRNDFLKEVIMRLKDPNI
IRLVGVQDDPLCMITDYMGNDLQFLSARLENKATQGLSDGESQGGTISYPM
LHVLGAGIASCGRYLAFLNEVHRDLATRNCLVGENFTIKIADFGMSRNLVAGDYRVQ
GLAVLPTRMAWECILMKFTTASDYWAFGVLWEVLMLCRSOPFQGLDQVIEVAG
EFFPDQGRQVYLSRPAPQTLVELMLRCWSREPRPPAQLHRLDADLNTV"
```

```
misc_feature 194..658
/feature="Nep"
/feature="encodes discoidin-like domain"
1340..1411
/feature="Nep"
/feature="encodes transmembrane domain"
1922..2809
/feature="Nep"
/feature="encodes tyrosine kinase domain"
BASE COUNT 724 a 1095 c 1075 g 765 t
ORIGIN
Query Match 74.8%; Score 895.6; DB 12; Length 3659;
Best Local Similarity 85.5%; Pred. No. 6.3e-184;
Matches:1026; Conservative 0; Mismatches 159; Indels 15; Gaps 2;
Qy 1 GATGCTGACATGAGGGACATTTGATCCTCCCAAGTGCCTATGCTGGGATCG 60
Db 158 GATGCTGACATGAGGGACATTTGATCCTCCCAAGTGCCTATGCTGGGATCG 217
Qy 61 GACCGGACATCCAGAGAGTACATCTCTCTCCAGCTCCTGTGTCAGATTCAC 120
Db 218 GACCGGACATCCAGAGAGTACATCTCTCTCCAGCTCCTGTGTCAGACTT 277
Qy 121 GCCCGCCACAGAGTTGGAGAGAGTACCGGGATGGGGCTGGTCCCGCAGGG 180
Db 278 GCCCGCCACAGAGTTGGAGAGAGTACCGGGATGGGGCTGGTCCCGCAGGG 337
Qy 181 GGTGTTCCCAAGGAGGAGTACTTGCAGTGGATCTTACAACTGCACCTGCTG 240
Db 338 GGTGTTCCCAAGGAGGAGTACTTGCAGTGGATCTTACAACTGCACCTGCTG 397
Qy 241 CTGGTGGGACCCAGGAGCGGATGCCGGGGCCCTGGGCAAGAGTCTTCCCG 300
Db 398 CTGGTGGGACCCAGGAGCGGATGCCGGGGCCCTGGGCAAGAGTCTTCCCG 457
Qy 301 CGGCTGCGTTACTCCCGGATGGTCCCGCTGGATGGCTGGAAGGACCTGGG 360
Db 458 CGGCTGCGTTACTCCCGGATGGTCCCGCTGGATGGCTGGAAGGACCTGGG 517
Qy 361 GAGGTGATCTCAGGCAATGAGGACCTCAGGAGTGGTGGTGAAGGACCTTGG 420
Db 518 GAGGTGATCTCAGGCAATGAGGACCTCAGGAGTGGTGGTGAAGGACCTTGG 577
Qy 421 ATGGTGGCCGACTGGTTCGCTTCTTACCCCGGGCTGACCGGGTATGAGCT 480
Db 578 ATGGTGGCCGACTGGTTCGCTTCTTACCCCGGGCTGACCGGGTATGAGCT 637
Qy 481 CGGCTAGAGCTCTATGGCTGCTCTGGAGGATGGACTCTCTTACACCGCC 540
Db 638 CGGCTAGAGCTCTATGGCTGCTCTGGAGGATGGACTCTCTTACATACAG 697
Qy 541 GGGGACAGCAATCTATTTATCTGAGG---CCGTGTACCTCAACGACTCCAC 597
Db 698 GGGGACAGCAATCTATTTATCTGAGG---CCGTGTACCTCAACGACTCCAC 757
Qy 598 CATACCTGGGCGGACTGCAGTATGGGGTCTGGGCCAGCTGGCAGATGGTGT 657
Db 758 TATACCTGGGCGGACTGCAGTATGGGGTCTGGGCCAGCTGGCAGATGGTGT 817
Qy 658 CTGGATGACTTTAGGAAGAGTACAGAGCTGGGGTCTGGCCAGGCTATGACT 717
Db 818 CTGGATGACTTTAGGAAGAGTACAGAGCTGGGGTCTGGCCAGGCTATGACT 877
Qy 718 TGAGCAACACAGCTTCTCCAGTGGCTATGTGAGATGGAGTTTGTGACCG 777
Db 878 TGAGCAACACAGCTTCTCCAGTGGCTATGTGAGATGGAGTTTGTGACCG 937
Qy 778 AGGGCCTTCCAGGCTATGAGTCCACTGTAAACAATGACACAGCTGGGAGCC 837
Db 938 AGAACCTTCCAGACCATGAGGTCCACTGTAAACAATGACACAGCTGGGAG 997
```


QY 781 GCCTTCAGGCTATGAGGTCAGGTCAGTAAACAACATCCACACAGCTGGAGCCGCTGTGCT 840
 Db 1015 TCTTTCAGACCATGAGGTCAGGTCAGTAAACAACATCCACACAGCTGGAGCCGCTGTGCT 1074
 QY 841 GCGGGGTGAATGTCGCTTCCGCGCTGGCCCTGCCATGCCCTGGAGGGGAGCCGCTG 900
 Db 1075 GCGGGGTGAATGTCGCTTAAAGGGTCCGGCCATGCCCTGGAGGGGAGGCTGTA 1134
 QY 901 CGCCACAACCTAGGGGCAACCTGGGGGAGCCCGAGAGCCCGGCTGTCTCAGTGGCCCTT 960
 Db 1135 CACCATGCCCTGGGGGTAGCTTGGAGAGCCCGAGAGCCCGGCTGTCTCAGTGGCCCTG 1194
 QY 961 GCGGGCGGTGCTGCTTCTGCAAGTCCGCTTCTTTCGCGGGCCCTGTTACTC 1020
 Db 1195 GGTGGTACAGTGGGCGCTTCTGCAAGTCCGCTTCTTTCGAGGACCTGGTACTC 1254
 QY 1021 TTCAGGAAATCTCTTCTATCTCTGATGTTGGTGAACAATCTCTTCCGCACTGGGAGGC 1080
 Db 1255 TTCAGTGAGATCTTCTTCTATCTCGATGTTGGTGAAGACTCTGCT-----GAC 1302
 QY 1081 ACCTTCCGCGCCAGCCCTTGTGGCGCTGGCCCACTCCCAACCTTCAGCAGCTTG 1140
 Db 1303 ACCTTCCCAAGCCCTTGTGGCGCTGGCCCACTCCCAACCTTCAGCAGCTTG 1362
 QY 1141 GAGCTGAGCCAGAGCCAGCAGCCGCTGGCCCAAGCCCGAGGAGCCCGCC 1197
 Db 1363 GAGCTGAGCCCGGGGTCAACAGCCTGTGGCCCAAGCCCGAGGAGCCCGCCAACTGCC 1419

RESULT 13
 HSU48705
 LOCUS
 DEFINITION Human receptor tyrosine kinase DDR gene, complete cds.
 ACCESSION U48705
 NID 91753221
 VERSION U48705.1 GI:1753221
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 12010)
 Sakuma S., Saya, H., Tada, M., Nakao, M., Fujiwara, T., Roth, J.A.,
 Sawamura, Y., Shinohe, Y. and Abe, H.
 Receptor protein tyrosine kinase DDR is up-regulated by p53 protein
 FEBS Lett. 398 (2-3), 165-169 (1996)
 97131588
 2 (bases 1 to 12010)
 Sakuma S., Tada, M., Saya, H., Sawamura, Y., Shinohe, Y. and Abe, H.
 Direct Submission
 Submitted (06-FEB-1996) Shiro Sakuma, Neurosurgery, Hokkaido
 University, North 15, West 7, Kita-ku, Sapporo 060, Japan
 Location/Qualifiers
 1. 12010
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 join(1856..2384,2576..2952,3213..3309,3841..4030,
 4138..4384,4910..5005,5114..5265,6351..6516,7250..7360,
 8467..8711,8860..8987,9226..9444,9920..10154,10747..10896,
 11015..12010)
 /product="tyrosine kinase DDR"
 1856..2384
 /number=1
 join(2197..2384,2576..2952,3213..3309,3841..4030,
 4138..4384,4910..5005,5114..5265,6351..6516,7250..7360,
 8467..8711,8860..8987,9226..9444,9920..10154,10747..10896,
 11015..11155)
 /codon_start=1
 /product="tyrosine kinase DDR"
 /protein_id="AAC50917.1"
 /db_xref="PID:g1754536"
 /db_xref="GI:1754536"
 /translation="MGPEALSSLLLLLVASGDADMKGHFDPAKCRYALGMDRTIPD

SDTASSWSSTAAHRSRLSSDGDGAWCPAGSVFPKEEYLVQDLQRHLVALVGT
 QGRHAGGLAKESRSYRLAYSRDGRWGMKDRWQEVISGNEDEGVVYLKDLGPPWV
 ARLVRYPRADRVMSVCLVELYGLWRDGLLSYAPVGTWYLSAENVYLNSTYDGH
 TVGLOYGGLQLADGVGLDDEPKSQELRVPGDYVGSNHSSSGIVVEFEFDR
 LRAFQWQVHCNNMHTLGLRPGVECFRRGPMAMWEGEPMRHNLGNGDPRARAV
 SVPLQGRVARFLQCFELFAGPMLLSEISFISDVNNSSPALGGTFPPAPWPPGPP
 TNFSLELEPRGQPVAKAEGSPTAILIGLVAIILLLIIIALMLRHLRRLLSKA
 ERYVEELTTHLSVPGDTLILNNRPGPREPPYQPRPNPSPHAPCVNPGSALL
 SNPARYLLATVARPPRGPPPTPAWAKPTQAYSGDYNEPEKPCAPILLPPONS
 PHYAADIYTIAGVTGNTYAVPALPPGAVGDPVPRDFRNLRFKELKGGQGEV
 HICEVSPQDLVSLDFLNVRKHPLLVAVKILRPDPRATNKNDFLKEVKIMSLKDP
 NIIRLGVCVQDDPLCMITDYMENGLNQLSAHQLEDAAGAGDGAAGQATISY
 PMLHLVAAQIASGRYLATLNFVRDLATRNCLVGENFTIKIADFGMSRLYAGDYR
 VQGRALVPIRWMAWECILMGKFTTASDYWAFVTVMEVLMCLRAOPFGOLTDEQVIE
 AGFEFRDQGRQVYLSRPPACPGLYELMLRCSRESEQRPPFSQLHRLFAEDALATV
 2576..2952
 /number=2
 3213..3309
 /number=3
 3841..4030
 /number=4
 4138..4384
 /number=5
 4910..5005
 /number=6
 5114..5265
 /number=7
 6351..6516
 /number=8
 7250..7360
 /number=9
 8467..8711
 /number=10
 8860..8987
 /number=11
 9226..9444
 /number=12
 9920..10154
 /number=13
 10747..10896
 /number=14
 11015..12010
 /number=15

BASE COUNT 2471 a 3250 c 3271 g 3006 t 12 others
 ORIGIN

Query Match 31.6%; Score 378.4; DB 11; Length 12010;
 Best Local Similarity 97.2%; Pred. No. 1.7e-72;
 Matches 385; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 119 CCGCCGCCACACAGGTTGAGAGCAGTACCGGGATGGGCGCTGGTCCCGCAGGCT 178
 Db 2560 CCCTTTCACCTTCAGTTGGAGAGCAGTACCGGGATGGGCGCTGGTCCCGCAGGCT 2619
 QY 179 CGGTGTTTCCCAAGGAGGAGGAGTACTTTCAGGTGGATCTACACAGCTCACCTGGTGG 238
 Db 2620 CGGTGTTTCCCAAGGAGGAGGAGTACTTTCAGGTGGATCTACACAGCTCACCTGGTGG 2679
 QY 239 CTCTGTGGGACCCAGGAGCAGTACCGGGGCGCTGGGCAAGGAGTCTTCCCGGAGCT 298
 Db 2680 CTCTGTGGGACCCAGGAGCAGTACCGGGGCGCTGGGCAAGGAGTCTTCCCGGAGCT 2739
 QY 299 ACCGGCTGCTTACTCCCGGATGGTCCCGCTGGATGGCTGGAAGACCGCTGGGGTC 358
 Db 2740 ACCGGCTGCTTACTCCCGGATGGTCCCGCTGGATGGCTGGAAGACCGCTGGGGTC 2799
 QY 359 AGGAGTGATCTCAGGCAATGAGGACCTCAGGAGTGGTGTCTGAAGACCTTGGGCCCC 418
 Db 2800 AGGAGTGATCTCAGGCAATGAGGACCTCAGGAGTGGTGTCTGAAGACCTTGGGCCCC 2859
 QY 419 CCATGTTGCCGACGAGTGTCTTACCCCGGCTGACCGGGTCTATGAGGCTCTGTC 478

Qy	13	AAGGACATTTTGATCCTGCCAAGTGCCTCGCTATGGCTTGGGATGCAGGACCCGGAACATC	72
Db	433	AAASCTCAGGTTAAATCAAGCTATATGCCGTATCTCTGGCATTCAGSAGGCCAGATT	492
Qy	73	CCAGACAGTCACATCTCTGCTCCAGCTCCTCGTCCAGATTCACATGCGCGCCGCACAGC	132
Db	493	CCAGATGAGGACATCACAGCTTCCAGTCAGTGGTCAGAGTCCACAGTGCCTCAAAATATGGA	552
Qy	133	AGGTTGAGAGCAGTAGTCACGGGATGGGCGCTTGGTCCCGCCGACGGTCTGGTTTTCCC AAG	192
Db	553	AGGCTGGACTCACAAGAAGGGATGAGGCTTGGTCCCTGAGATTCAGATGGAAACCTGAT	612
Qy	193	GA---GGAGAGTACTTGCAGGTGGATCTACAAACAGCTGCACCTGGTGGCTCTGGTGGCG	249
Db	613	GACCTGAAGGAGTTCTGTCAGATGTACTTGCACACCCCTCCATTTTATCACTCTGTTGGGG	672
Qy	250	ACCAGGAGCGGATCGCGGGGCTTGGCGAAGGAGTTCTCCCGGAGCTACCGGCTCGGT	309
Db	673	ACCAGGGGGCCCATCGAGAGGTATGCGATCGAGTTTGGCCCCCATGTCAAGATCAAT	732
Qy	310	TACTCCCGGATGGTCCCGCTGGATGGGCTGGAAGACCGCTGGGGTCAGGAGGTGATC	369
Db	733	TACAGTCGGATGSCACTCGCTGGATCTTGGCGGAACCGTCATGGGAACAGTGGCTG	792
Qy	370	TCAGGCAATGAGGACCCCTGAGGAGTGGTGTGAAGACCTTGGGCGCCCGCATGGTTGCC	429
Db	793	GATGGAATAAGTAACCCCTATGACATTTTCCCTAAAGGAGCTTGGAGCCGCCCATTTGAGCC	852
Qy	430	CGACTGGTTCGCTTACCCCGGGCTGACCGGGTCATGAGGCTCTGCTCGGGGTAGAG	489
Db	853	AGATTTTCGGGTTTCAATCCAGTCACCGACCACTCCATGAATGTGTGATGAGAGTGGAG	912
Qy	490	CTCTATGGCTCCCTCTGGAGGATGACCTCTGTCTTACACGCCCGCTGTGGGGCAGACA	549
Db	913	CTTTACGGCTGTCTCTGGCTAGATGCTTGGTGTCTTACAATGCTCCAGCTGGGCACAG	972
Qy	550	ATGTAATTTATCTGAGG-----CGTGTACCTCAAGGACTCCAACCTATGACGGACATACC	603
Db	973	TTTGACTCCCTCGAGGTTCCATCATTTATCTCAATGATTTCTGTCTATGATGAGAGCTGTT	1032

Db 2605 AGGCTGACTCAGAGAGGGGATGGAGCCTGGTGGCCCTGAGATTCAGTGGAACTGAT 2546
Qy 193 GA---GGAGGAGTACTTGCAGGTGGATACAACTGACCTGGTGGCTCTGGTGGC 249
Db 2545 GACCTGAAGGAGTTCTGCAGATTGACTTGCACACCCCTCCATTTATCACTCTGGTGGG 2486
Qy 250 ACCAGGACCGCATGCGGGGGCTGGGAGAGGATTCCTCCGGAGCTACCGGCTGCGT 309
Db 2485 ACCAGGGGGCCATGAGGAGGTCATGGCATCGAGTTGCCCCCATGTACAAGATCAAT 2426
Qy 310 TACTCCCGGATGCTGCGCGCTGGATGGGTGGAAGGACCGCTGGGGTCAGGAGGTGATC 369
Db 2425 TACAGTCGGATGACATCGCTGGATCTCTTGGGGAACCGTATGGGAACAGGTGCTG 2366
Qy 370 TCAGGCAATGAGCCCTGAGGGAGTGGTCTGAAGGACCTTGGGCCCCCATGGTTGCC 429
Db 2365 GATGAAATAGTAAACCTATGACATTTCTAAAGGACCTGGAGCCGCCATTTAGCC 2306
Qy 430 CGACTGGTTCCTTCTACCCCGGGCTGACCGGGTCAAGCGTCTCTGCGGGTAGAG 489
Db 2305 AGATTGTGCGGTTTCATCCAGTCACCGACCATCCCATGAATGTGTATGAGAGTGGAG 2246
Qy 490 CTCATGCTGCTGCTGAGGAGTGGACTCTCTGTCTTACACCGCCCTGTGGGCGACACA 549
Db 2245 CTTACGGCTGTCTGCTAGATGGCTTGGTGTCTTACAATGCTCCAGCTGGGCGACAG 2186
Qy 550 ATGATTTATCTGAGGC-----CGTGTACTCAACGACTCCACCTATGACGGACATACC 603
Db 2185 TTTGTACTCCCTGGAGGTTCCATCATTTATCTGAATGATTCGTCTATGATGAGCTGTT 2126
Qy 604 GTGGCGGACTGACAGTATGGGGGTCTGGGCCAGCTGGCAGATGGTGTGGGGCTGGAT 663
Db 2125 GGATACAGCATGACAGAAGG---CTAGGCCAATTGACCGATGGTGTCTGCGCTGGAC 2069
Qy 664 GACTTTAGGAGAGTCAAGGACTCGGGTCTGGCGAGCTATGACTATGTGGGATGGAGC 723
Db 2068 GATTTACCCAGACCCATGAATACACGCTGGCGCGGCTATGACTATGTGGGCTGGCGG 2009
Qy 724 AACACAGCTTCTCCAGTGGCTATGTGGAGATGGAGTTTGTAGTTGACCGCTGAGGGCC 783
Db 2008 AACGAGATGCCACCAATGGCTACATTTGAGATCATGTTGNAATTGACCCGATCAGGAAT 1949
Qy 784 TTCAGGCTATGACAGGTCCACTGTAAACAATGACACAGCTGGAGGCCCTCTGCCCTGGC 843
Db 1948 TTCATACCATGAAGTCCACTGCAACAACATGTTGCTAAAGGTGTGAAGATCTTTAAG 1889
Qy 844 GGGTGGATGCTGCTTCCGCGCTGGCCCTGCCATGGCTGGAGGGGGAGCCCATGCCG 903
Db 1888 GAGGTACAGTGTCTACTTCCGCTCTGAA---GCCAGTGAAGTGGAAACCTAATGCCATTTC 1832
Qy 904 CACAACTAGGGGGCAACCTGGGGGACCCAGAGCCGGGTGCTCAGTGGCCCTTGGC 963
Db 1831 TTCCCTTGTCTGGATGAGCTCAACCCAGTGTCTGGTTGTACGGTGGCTCTCCAC 1772
Qy 964 GGCGGTGGGTGCGTCTTCTCAGTGGCGCTTCTCTTGGGGGGCCCTGGTTACTCTTC 1023
Db 1771 CACCGAATGGCCAGTGGCCATCAAGTGTCAATACCATTTTGCAGATACCTGGATGATTC 1712
Qy 1024 AGCGAATCTCCTCACTCTGATGTGGTGAACAATTCCTCCGGGACCTGGGAGGCACC 1083
Db 1711 AGTGAGATCACCTTCCATCAGATGCTGCAATGTAAACAACACTCTGAAGCCCTGCCACC 1652
Qy 1084 TTCGCCGACGCCCC 1098
Db 1651 TCTCTATGGCACCC 1637

Search completed: November 4, 1999, 07:31:51
Job time: 6510 sec

